BEST AVAILABLE COPY

METHODS OF IDENTIFYING OPTIMAL VARIANTS OF PEPTIDE EPITOPES

BACKGROUND OF THE INVENTION

Field of the Invention

[0001] This invention relates to the field of biology. In a particular embodiment, it relates to peptides, polynucleotides, and compositions useful to monitor or elicit an immune response to selected antigens, and methods of identifying such peptides and polynucleotides.

Related Art

[0002] HLA class I molecules are expressed on the surface of almost all nucleated cells. Following intracellular processing of antigens, epitopes from the antigens are presented as a complex with the HLA class I molecules on the surface of such cells. CTL recognize the peptide-HLA class I complex, which then results in the destruction of the cell bearing the HLA-peptide complex directly by the CTL and/or via the activation of non-destructive mechanisms *e.g.*, the production of interferon, that inhibit viral replication.

[0003] **Human Immunodeficiency Virus.** Acquired immunodeficiency syndrome (AIDS) caused by infection with human immunodeficiency virus-1 (HIV-1) represents a major world health problem. Estimates indicate that about 16,000 people worldwide are infected with HIV each day.

[0004] The development of anti-viral drugs has been a major advancement in reducing viral loads in HIV infected patients. Highly active retroviral therapy (HAART) has been shown to reduce viremia to nearly undetectable levels. However, current drug therapies are not practicable as a long term solution to the HIV epidemic. HAART therapy is severely limited due to poor tolerance for the drugs and the emergence of drug-resistant virus. Moreover, replication competent HIV persists in the lymphoid tissue of patients who have responded to HAART, thus serving as a reservoir of virus. Lastly, current anti-retroviral drug therapies have little impact upon the global epidemic: almost 90% of the world's HIV infected population resides within countries lacking financial resources for these drugs. Thus, a need exists for an efficacious vaccine to both prevent and treat HIV infection.

[0005] Virus-specific, human leukocyte antigen (HLA) class I-restricted cytotoxic T lymphocytes (CTL) are known to play a major role in the prevention and clearance of virus infections in vivo (Oldstone *et al.*, Nature 321:239, 1989; Jamieson et al., J. Virol. 61:3930, 1987; Yap et al, Nature 273:238, 1978; Lukacher et al., J. Exp. Med. 160:814, 1994; McMichael et al., N. Engl. J. Med. 309:13, 1983; Sethi et al., J. Gen. Virol. 64:443, 1983; Watari et al., J. Exp. Med. 165:459, 1987; Yasukawa et al., J. Immunol. 143:2051, 1989; Tigges et al., J. Virol. 66:1622, 1993; Reddenhase et al., J. Virol. 55:263, 1985; Quinnan et al., N. Engl. J. Med. 307:6, 1982).

[0006] While immune correlates of protective immunity against HIV infection are not well defined, there is a growing body of evidence that suggests CTL are important in controlling HIV infection. HIV-specific CTL responses can be detected early in infection and the appearance of the responses corresponds to the time in infection at which initial viremia is reduced (Pantaleo *et al., Nature* 370:463, 1994; Walker *et al., Proc. Natl. Acad. Sci.* 86:9514, 1989). In addition, HIV replication in infected lymphocytes can be inhibited by incubation with autologous CTL (*see, e.g.*, Tsubota *et al., J. Exp. Med.* 169:1421, 1989). These data are supported by recent studies that indicate CTL are required for controlling viral replication in a SIV/rhesus animal model (Schmitz *et al., Science* 283:857, 1999), and additionally supported by studies that demonstrate that CTL exert selective pressure on HIV populations as evidenced by the eventual predominance of viruses with amino acid replacements in those regions of the virus to which CTL responses are directed (*see, e.g.*, Borrow *et al., Nature Med.* 3:205-211, 1997; Price *et al., Proc. Nat. Acad. Sci.* 94:12890-1895, 1997; Koenig *et al., Nature Med.* 1:330-336, 1995; and Haas *et al., J. Immunol.* 157:4212-4221, 1996).

[0007] Virus-specific T helper lymphocytes are also known to be critical for maintaining effective immunity in chronic viral infections. Historically, HTL responses were viewed as primarily supporting the expansion of specific CTL and B cell populations; however, more recent data indicate that HTL may directly contribute to the control of virus replication. For example, a decline in $CD4^+$ T cells and a corresponding loss in HTL function characterize infection with HIV (Lane *et al., New Engl. J. Med.* 313:79, 1985). Furthermore, studies in HIV infected patients have also shown that there is an inverse relationship between virus-specific HTL responses and viral load, suggesting that HTL play a role in viremia (*see, e.g.*, Rosenberg *et al., Science* 278:1447, 1997).

[0008] A fundamental challenge in the development of an efficacious HIV vaccine is the heterogeneity observed in HIV. The virus, like some other infectious agents including retroviruses, rapidly mutates during replication resulting in the generation of virus that can escape anti-viral therapy and immune recognition (Borrow et al., Nature Med. 3:205, 1997). In addition, HIV can be classified into a variety of subtypes that exhibit significant sequence divergence (*see, e.g.*, Lukashov *et al., AIDS* 12:S43, 1998). In view of the heterogeneous nature of HIV, and the

heterogeneous immune response observed with HIV infection, induction of a multi-specific cellular immune response directed simultaneously against multiple HIV epitopes appears to be important for the development of an efficacious vaccine against HIV. There is a need to establish such vaccine embodiments which elicit immune responses of sufficient breadth and vigor to prevent and/or clear HIV infection.

[0009] **Hepatitis B Virus.** Chronic infection by hepatitis B virus (HBV) affects at least 5% of the world's population and is a major cause of cirrhosis and hepatocellular carcinoma (Hoofnagle, J., *N. Engl. J. Med.* 323:337, 1990; Fields, B. and Knipe, D., In: *Fields Virology* 2:2137, 1990). The World Health Organization lists hepatitis B as a leading cause of death worldwide, close behind chronic pulmonary disease, and more prevalent than AIDS. Chronic HBV infection can range from an asymptomatic carrier state to continuous hepatocellular necrosis and inflammation, and can lead to hepatocellular carcinoma.

[0010] The immune response to HBV is believed to play an important role in controlling hepatitis B infection. A variety of humoral and cellular responses to different regions of the HBV nucleocapsid core and surface antigens have been identified. T cell mediated immunity, particularly involving class I human leukocyte antigen-restricted cytotoxic T lymphocytes (CTL), is believed to be crucial in combatting established HBV infection.

[0011] Several studies have emphasized the association between self-limiting acute hepatitis and multispecific CTL responses (Penna, A. et al., *J. Exp. Med.* 174:1565, 1991; Nayersina, R. et al., *J. Immunol.* 150:4659, 1993). Spontaneous and interferon-related clearance of chronic HBV infection is also associated with the resurgence of a vigorous CTL response (Guidotti, L. G. et al., *Proc. Natl. Acad. Sci. USA* 91:3764, 1994). In all such cases the CTL responses are polyclonal, and specific for multiple viral proteins including the HBV envelope, core and polymerase antigens. By contrast, in patients with chronic hepatitis, the CTL activity is usually absent or weak, and antigenically restricted.

[0012] The crucial role of CTL in resolution of HBV infection has been further underscored by studies using HBV transgenic mice. Adoptive transfer of HBV-specific CTL into mice transgenic for the HBV genome resulted in suppression of virus replication. This effect was primarily mediated by a non-lytic, lymphokine-based mechanism (Guidotti, L. G. et al., *Proc. Natl. Acad. Sci. USA* 91:3764, 1994; Guidotti, L.

G., Guilhot, S., and Chisari, F. V. *J. Virol.* 68:1265, 1994; Guidotti, L. G. et al., *J. Virol.* 69:6158, 1995; Gilles, P. N., Fey, G., and Chisari, F. V., *J. Virol.* 66:3955, 1992).

[0013] As is the case for HLA class I restricted responses, HLA class II restricted T cell responses are usually detected in patients with acute hepatitis, and are absent or weak in patients with chronic infection (Chisari, F. V. and Ferrari, C., *Annu. Rev. Immunol.* 13:29, 1995). HLA Class II responses are tied to activation of helper T cells (HTLs) Helper T lymphocytes, which recognize Class II HLA molecules, may directly contribute to the clearance of HBV infection through the secretion of cytokines which suppress viral replication (Franco, A. et al., *J. Immunol.* 159:2001, 1997). However, their primary role in disease resolution is believed to be mediated by inducing activation and expansion of virus-specific CTL and B cells.

[0014] In view of the heterogeneous immune response observed with HBV infection, induction of a multi-specific cellular immune response directed simultaneously against multiple epitopes appears to be important for the development of an efficacious vaccine against HBV. There is a need to establish vaccine embodiments that elicit immune responses that correspond to responses seen in patients that clear HBV infection. Epitope-based vaccines appear useful.

[0015] **Hepatitis C Virus.** Hepatitis C virus (HCV) infection is a global human health problem with approximately 150,000 new reported cases each year in the U.S. alone. HCV is a single stranded RNA virus, and is the etiological agent identified in most cases of non-A, non-B post-transfusion and post-transplant hepatitis, and is a common cause of acute sporadic hepatitis (Choo *et al.*, *Science* 244:359, 1989; Kuo *et al.*, *Science* 244:362, 1989; and Alter *et al.*, in: *Current Perspective in Hepatology*, p. 83, 1989). It is estimated that more than 50% of patients infected with HCV become chronically infected and, of those, 20% develop cirrhosis of the liver within 20 years (Davis *et al.*, *New Engl. J. Med.* 321:1501, 1989; Alter *et al.*, in: *Current Perspective in Hepatology*, p. 83, 1989; Alter *et al.*, *New Engl. J. Med.* 327:1899, 1992; and Dienstag, J. L. *Gastroenterology* 85:430, 1983). Moreover, the only therapy available for treatment of HCV infection is interferon-α. Most patients are unresponsive, however, and among the responders, there is a high recurrence rate within 6-12 months of cessation of treatment (Liang *et al.*, *J. Med. Virol.* 40:69, 1993). Ribaviron, a guanosine analog with a broad spectrum activity against many RNA and DNA viruses, has been shown in clinical trials to be effective against chronic

HCV infection when used in combination with interferon- α (*see, e.g.*, Poynard *et al.*, *Lancet* 352:1426-1432, 1998; Reichard *et al.*, *Lancet* 351:83-87, 1998) However, the response rate is still well below 50%.

[0016] Virus-specific, human leukocyte antigen (HLA) class I-restricted cytotoxic T lymphocytes (CTL) are known to play a major role in the prevention and clearance of virus infections *in vivo* (Oldstone *et al.*, *Nature* 321:239, 1989; Jamieson *et al.*, *J. Virol.* 61:3930, 1987; Yap *et al*, *Nature* 273:238, 1978; Lukacher *et al.*, *J. Exp. Med.* 160:814, 1994; McMichael *et al.*, *N. Engl. J. Med.* 309:13, 1983; Sethi *et al.*, *J. Gen. Virol.* 64:443, 1983; Watari *et al.*, *J. Exp. Med.* 165:459, 1987; Yasukawa *et al.*, *J. Immunol.* 143:2051, 1989; Tigges *et al.*, *J. Virol.* 66:1622, 1993; Reddenhase *et al.*, *J. Virol.* 55:263, 1985; Quinnan *et al.*, *N. Engl. J. Med.* 307:6, 1982).

[0017] In view of the heterogeneous immune response observed with HCV infection, induction of a multi-specific cellular immune response directed simultaneously against multiple HCV epitopes appears to be important for the development of an efficacious vaccine against HCV. There is a need, however, to establish vaccine embodiments that elicit immune responses that correspond to responses seen in patients that clear HCV infection.

[0018] **Human Papillomavirus.** Human papillomavirus (HPV) is a member of the papillomaviridae, a group of small DNA viruses that infect a variety of higher vertebrates. More than 80 types of HPVs have been identified. Of these, more than 30 can infect the genital tract. Some types, generally types 6 and 11, may cause genital warts, which are typically benign and rarely develop into cancer. Other strains of HPV, "cancer-associated", or "high-risk" types, can more frequently lead to the development of cancer. The primary mode of transmission of these strains of HPV is through sexual contact.

[0019] The main manifestations of the genital warts are cauliflower-like condylomata acuminata that usually involve moist surfaces; keratotic and smooth papular warts, usually on dry surfaces; and subclinical "flat" warts, which are found on any mucosal or cutaneous surface (Handsfield, H., *Am. J. Med.* 102(5A):16-20, 1997). These warts are typically benign but are a source of inter-individual spread of the virus (Ponten, J. & Guo, Z., *Cancer Surv.* 32:201-29, 1998). At least three HPV strains associated with genital warts have been identified: type 6a (see, *e.g.*, Hofmann, K.J., *et al., Virology 209(2):506-*518,

1995), type 6b (see, *e.g.*, Hofmann *et al.*, supra) and type 11 (see, *e.g.*, Dartmann, K. *et al.*, *Virology* 151(1):124-130, 1986).

[0020] Cancer-associated HPVs have been linked with cancer in both men and women; they include, but are not limited to, HPV-16, HPV-18, HPV-31, HPV-45, HPV-33 and HPV-56. Other HPV strains, including types 6 and 11 as well as others, *e.g.*, HPV-5 and HPV-8, are less frequently associated with cancer. The high risk types are typically associated with the development of cervical carcinoma and premalignant lesions of the cervix in women, but are also associated with similar malignant and premalignant lesions at other anatomic sites within the lower genital or anogenital tract. These lesions include neoplasia of the vagina, vulva, perineum, the penis, and the anus. HPV infection has also been associated with respiratory tract papillomas, and rarely, cancer, as well as abnormal growth or neoplasia in other epithelial tissues. *See, e.g.* VIROLOGY, 2ND ED, Fields *et al.*, Eds. Raven Press, New York, 1990, Chapters 58 and 59, for a review of HPV association with cancer.

[0021] The HPV genome consists of three functional regions, the early region, the late region, and the "long control region". The early region gene products control viral replication, transcription and cellular transformation They include the HPV E1 and E2 proteins, which play a role in HPV DNA replication, and the E6 and E7 oncoproteins, which are involved in the control of cellular proliferation. The late region include the genes that encode the structural proteins L1 and L2, which are the major and minor capsid proteins, respectively. The "long control region" contains such sequences as enhancer and promoter regulatory regions.

[0022] HPV expresses different proteins at different stages of the infection, for example early, as well as late, proteins. Even in latent infections, however, early proteins are often expressed and are therefore useful targets for vaccine-based therapies. For example, high-grade dysplasia and cervical squamous cell carcinoma continue to express E6 and E7, which therefore can be targeted to treat disease at both early and late stages of infection.

[0023] Treatment for HPV infection is often unsatisfactory because of persistence of virus after treatment and recurrence of clinically apparent disease is common. The treatment may require frequent visits to clinics and is not directed at elimination of the virus but at clearing warts. Because of persistence of virus after treatment, recurrence of clinically apparent disease is common.

[0024] Thus, a need exists for an efficacious vaccine to both prevent and treat HPV infection and to treat cancer that is associated with HPV infection. Effective HPV vaccines would be a significant advance in the control of sexually transmissable infections and could also protect against clinical disease, particularly cancers such as cervical cancer. (*see, e.g.*, Rowen, P. & Lacey, C., *Dermatologic Clinics* 16(4):835-838, 1998).

[0025] Virus-specific, human leukocyte antigen (HLA) class I-restricted cytotoxic T lymphocytes (CTL) are known to play a major role in the prevention and clearance of virus infections *in vivo* (Oldstone *et al.*, *Nature* 321:239, 1989; Jamieson *et al.*, *J. Virol.* 61:3930, 1987; Yap *et al*, *Nature* 273:238, 1978; Lukacher *et al.*, *J. Exp. Med.* 160:814, 1994; McMichael *et al.*, *N. Engl. J. Med.* 309:13, 1983; Sethi *et al.*, *J. Gen. Virol.* 64:443, 1983; Watari *et al.*, *J. Exp. Med.* 165:459, 1987; Yasukawa *et al.*, *J. Immunol.* 143:2051, 1989; Tigges *et al.*, *J. Virol.* 66:1622, 1993; Reddenhase *et al.*, *J. Virol.* 55:263, 1985; Quinnan *et al.*, *N. Engl. J. Med.* 307:6, 1982).

[0026] Virus-specific T helper lymphocytes are also known to be critical for maintaining effective immunity in chronic viral infections. Historically, HTL responses were viewed as primarily supporting the expansion of specific CTL and B cell populations; however, more recent data indicate that HTL may directly contribute to the control of virus replication. For example, a decline in $CD4^+$ T cells and a corresponding loss in HTL function characterize infection with HIV (Lane *et al.*, *New Engl. J. Med.* 313:79, 1985). Furthermore, studies in HIV infected patients have also shown that there is an inverse relationship between virus-specific HTL responses and viral load, suggesting that HTL plays a role in viremia (*see, e.g.*, Rosenberg *et al.*, *Science* 278:1447, 1997).

[0027] The development of vaccines with prophylactic and therapeutic efficacy against HPV is ongoing. Early vaccine development was hampered by the inability to culture HPV. With the introduction of cloning techniques and protein expression, however, some attempts have been made to stimulate humoral and CTL response to HPV (*See, e.g.*, Rowen, P. & Lacey, C., *Dermatologic Clinics* 16(4):835-838 (1998)). Studies to date, however, have been inconclusive.

[0028] Activation of T helper cells and cytotoxic lymphocytes (CTLs) in the development of vaccines has also been analyzed. Lehtinen, M., *et al.* for instance, has shown that some peptides from the E2 protein of HPV type 16 activate T helper cells and CTLs (*Biochem. Biophys. Res. Commun.* 209(2):541-6 (1995). Similarly, Tarpey *et al*, has shown that some peptides from HPV type 11 E7 protein can stimulate human HPV-specific CTLs *in*

vitro (*Immunology* 81:222-227 (1994)) and Borysiewicz *et al.* have reported a recombinant vaccinia virus expressing HPV 16 and HPV 17 E6 and E7 that stimulated CTL responses in at least one patient (*Lancet* 347:1347-1357, 1996).

[0029] ***Plasmodium falciparum* and Malaria.** Malaria, which is caused by infection with the parasite *Plasmodium falciparum* (PF), represents a major world health problem. Approximately 500 million people in the world are at risk from the disease, with approximately 200 million people actually harboring the parasites. An estimated 1 to 2 million deaths occur each year due to malaria. (Miller *et al.*, *Science* 234:1349, 1986).

[0030] Fatal outcomes are not confined to first infections, and constant exposure is apparently a prerequisite for maintaining immunity. Naturally acquired sterile immunity is rare, if it exists at all. Accordingly, major efforts to develop an efficacious malaria vaccine have been undertaken.

[0031] Human volunteers injected with irradiated PF sporozoites are resistant to subsequent sporozoite challenges, which demonstrates that development of a malaria vaccine is indeed immunologically feasible. Furthermore, these immune individuals developed a vigorous response, including antibodies, and cytotoxic T lymphocyte (CTL) and helper T lymphocyte (HTL) components, directed against multiple antigens. Reproducing the breadth and multiplicity of this response in a vaccine, however, is a task of large proportions. The epitope approach, as described herein, may represent a solution to this challenge, in that it allows the incorporation of various antibody, CTL and HTL epitopes, from various proteins, in a single vaccine composition.

[0032] Anti-sporozoite antibodies are by themselves, in general, not completely efficacious in clearing the infection (Egan *et al.*, *Science* 236:453, 1987). However, high concentrations of antibodies directed against the repeated region of the major B cell antigen of the sporozoite/circumsporozoite protein (CSP) have been shown to prevent liver cell infection in certain experimental models (Egan *et al.*, *Science* 236:453, 1987; Potocnjak, P. *et al.*, *Science* 207:71, 1980). The present inventors have shown that constructs encompassing CSP-repeat B cell epitopes and the optimized helper epitope PADRE™ (San Diego, CA) are highly immunogenic, and can protect *in vitro* against sporozoite invasion in both mouse and human liver cells, and protect mice *in vivo* against live sporozoite challenge (Franke *et al.*, *Vaccine* 17:1201-1205, 1999)

[0033] PF-specific $CD4^+$ T cells also have a role in malarial immunity beyond providing help for B cell and CTL responses. Experiments by Renia *et al.* (Renia, *et al.*, *Proc. Natl. Acad. Sci. USA* 88:7963, 1991) demonstrated that HTLs directed against the *Plasmodium yoelli* CS protein could in fact adoptivley transfer protection against malaria.

[0034] Considerable data implicate CTLs in protection against pre-erythrocytic-stage malaria. $CD8^+$ CTLs can eliminate *Plasmodium berghei-* or *Plasmodium yoelii*-infected mouse hepatocytes from in vitro culture in a major histocompatibility complex (MHC)-restricted and antigen-restricted manner (Hoffman *et al.*, *Science* 244:1078-1081, 1989; Weiss *et al.*, J. *Exp. Med.* 171:763-773, 1990). Further, it has also been shown that the immunity that developed in mice vaccinated with irradiated sporozoites is also dependent upon the present of CD8+ T cells. These T cells accumulate in inflammatory liver infiltrates subsequent to challenge. Passive transfer of circumsporozoite (CSP)-specific CTL clones as long as three hours after inoculation of sporozoites (*i.e.*, after the parasites have left the bloodstream and infected liver cells) were capable of protecting animals against infection (Romero *et al.*, *Nature* 341:323, 1989).

[0035] It is notable that CTL-restricted responses directed against a single antigen are insufficient to protect mice with different MHC alleles, and a combination of multiple antigens was required even to protect mice from the most common laboratory strains of *Plasmodium*. These data indicate that a combination of epitopes form several antigens is necessary to elicit a protective CTL response.

[0036] Indirect evidence that CTLs are important in protective immunity against Pf in humans has also accumulated. It has been reported that cytotoxic $CD8^+$ T cells can be identified in humans immunized with PF sporozoites (Moreno, *et al.*, *Int. Immunol.* 3:997, 1991). Further, humans immunized with irradiated sporozoites or naturally exposed to malaria can generate a CTL response to the pre-erythrocytic-stage antigens, CSP, sporozoite surface protein 2 (SSP2), liver-stage antigen-1 (LSA-1), and exported protein-1 (Exp-1) (*see, e.g.* Malik *et al.*, *Proc. Natl. Acad. Sci. USA* 88, 3300-3304, 1991; Doolan *et al.*, *Int. Immunol.* 3:511-516, 1991; Hill *et al.*, *Nature* 360:434-439, 1992). Additionally, there is evidence that the polymorphism within the CSP may be the result of selection by CTLs of parasites that express variant forms (MCutchan and Water, Immunol. Lett. 25:23-26, 1990). This is based on the observation that the variation is nonsynonymous at the nucleotide level, thereby indicating selective pressure at the protein level. The polymorphism primarily maps to identified CTL and T helper epitopes (Doolan *et al.*, *Int.*

Immunol. 5:27-46, 1993); and CTL responses to some of the parasite variants do not cross-react (Hill *et al.*, *supra*). Finally, the MHC class I human leukocyte antigen (HLA)-Bw53 has been associated with resistance to severe malaria in The Gambia, and CTLs to a conserved epitope restricted by the HLA-Bw53 allele have been identified on *P. falciparum* LSA-1 (Hill *et al.*, *Nature* 352:595-600, 1991; Hill *et al.*, *Nature* 340:434-439, 1992). Since HLA-Bw53 is found in 15%-40% of the population of sub-Saharan Africa but in less than 1% of Caucasians and Asians, these data suggest evolutionary selection on the basis of protection against severe malaria.

[0037] Thus, antibody, and both HLA class I and class II restricted responses directed against multiple sporozoite antigens appear to be involved in generating protective immunity to malaria. Furthermore, several important antigenic epitopes against which humoral and cellular immunity is focused have already been exactly delineated.

[0038] In view of the heterogeneous immune response observed with PF infection, induction of a multi-specific cellular immune response directed simultaneously against multiple PF epitopes appears to be important for the development of an efficacious vaccine against PF. There is a need, however, to establish vaccine embodiments that elicit immune responses that correspond to responses seen in patients that clear PF infection.

[0039] **Epitope-Based Vaccines.** The use of epitope-based vaccines has several advantages over current vaccines. The epitopes for inclusion in such a vaccine are to be selected from conserved regions of viral or tumor-associated antigens, in order to reduce the likelihood of escape mutants. The advantage of an epitope-based approach over the use of whole antigens is that there is evidence that the immune response to whole antigens is directed largely toward variable regions of the antigen, allowing for immune escape due to mutations. Furthermore, immunosuppressive epitopes that may be present in whole antigens can be avoided with the use of epitope-based vaccines.

[0040] Additionally, with an epitope-based vaccine approach, there is an ability to combine selected epitopes (CTL and HTL) and additionally to modify the composition of the epitopes, achieving, for example, enhanced immunogenicity. Accordingly, the immune response can be modulated, as appropriate, for the target disease. Similar engineering of the response is not possible with traditional approaches.

[0041] Another major benefit of epitope-based immune-stimulating vaccines is their safety. The possible pathological side effects caused by infectious agents or whole protein antigens, which might have their own intrinsic biological activity, is eliminated.

[0042] An epitope-based vaccine also provides the ability to direct and focus an immune response to multiple selected antigens from the same pathogen. Thus, patient-by-patient variability in the immune response to a particular pathogen may be alleviated by inclusion of epitopes from multiple antigens from that pathogen in a vaccine composition. A "pathogen" may be an infectious agent or a tumor associated molecule.

[0043] One of the most formidable obstacles to the development of broadly efficacious epitope-based immunotherapeutics has been the extreme polymorphism of HLA molecules. In the past, effective non-genetically biased coverage of a population has been a task of considerable complexity; such coverage has required that epitopes be used specific for HLA molecules corresponding to each individual HLA allele. Therefore, impractically large numbers of epitopes would been required in order to cover ethnically diverse populations. Recently, methods have been developed that allow the identification of epitopes that bind multiple HLA molecules. Therefore, epitope-based vaccines can be designed that contain epitopes which, either individually or in combination, bind a greater number of HLA molecules. The resulting epitope-based vaccines have a greater breadth of population coverage across one or more continents and even worldwide.

[0044] **Variation in Epitopes of Infectious Agents.** A challenge in the development of effective vaccines against infectious agents such as hepatitis B virus (HBV) (47, 60) hepatitis C virus (HCV) (61-63), human papilloma virus (HPV) (64, 65) *Plasmodium falciparum* (66), and human immunodeficiency virus (HIV-1) is the protein sequence variation associated with different isolates. This variation is the result of gene sequence mutations. When such mutations occur in regions encoding epitopes recognized by cytotoxic T-lymphocytes (CTL), they provide a mechanism for escape of the agent from immune system control.

[0045] HIV-1 represents an infectious agent with an especially high frequency of sequence variation. The sequence variation associated with HIV-1 proteins from related isolates, members of the same clades or types, as well as unrelated isolates, is well documented (1). Viral escape from CTL induced as the result of natural infection or vaccines was documented in nonhuman primate models where the mechanism behind this

escape was mutation of the primary anchor residues in dominant CTL epitopes (5-9). Viral escape from HIV-specific CTL has also been strongly implied by data obtained from HIV-1 infected individuals whose disease status change, including the transition from acute to chronic infection (10, 11), loss of stable control of viral replication and subsequent progression to AIDS (4, 12) or mother-to-child transmission (13). Thus, HIV-1 genetic and protein sequence variation represent a significant challenge to immune system-based control of viral replication, both within infected individuals and within populations.

[0046] While the public health need for a vaccine against HIV-1 is well recognized and accepted, the genetic variation of HIV-1 isolates represents a highly significant obstacle (1, 14-16). Several strategies have been proposed, some of which include:

(1) Designing vaccines on HIV-1 types prevalent within small, well defined populations or geographical regions, such as individual countries or regions, and producing multiple different vaccines for exclusive use within these countries or regions (16).

(2) Use of HIV-1 ancestral or consensus sequences based on HIV types present in larger target populations, such as groups of neighboring countries or continents (15, 17-19).

(3) Incorporation of viral gene products obtained from multiple different virus isolates, representing diversely different types or clades, into a single 'multi-valent' vaccine.

[0047] Related vaccine design concepts that incorporate many of the advantages associated with the approaches described above are the use of highly conserved regions or epitopes derived from these regions as the basis of the vaccine. The logic behind this approach is that conserved regions of the viral genome are those that have been maintained through the evolution of HIV-1 because changes impact gene product function and general viral fitness. This theory is consistent with analyses of HIV-1 protein sequence data which demonstrated that CTL epitopes are concentrated in conserved regions and that regions devoid of CTL epitopes are the most variable (20). Additional support comes from published reports describing CTL responses, induced as the result of

natural infection or vaccination, that recognize viral proteins or epitopes common to viral isolates from diverse types or clades (21-26). Broad function CTL responses are also known to be correlated with slower progression to AIDS, at least for certain carefully studied populations (27, 28). Despite these reports and the clustering of CTL epitopes in conserved regions of HIV-1 gene products, amino acid sequence variation of analogous regions and epitopes from different viral isolates, both within the same type or clade and from different types, remains significant. There are currently no rules guiding the selection of conserved regions of CTL epitopes for use in vaccines other than the use of amino acid sequence identity (29).

[0048] A clear understanding of how CTL recognize pathogen infected cells has emerged over the past decade. It is now well established that small fragments of pathogen-derived proteins are generated, defined as peptide epitopes generally 8-11 amino acids in length, which bind to HLA-A, -B, or -C (human Class I Major Histocompatability Molecules) molecules expressed on the cell surface. Sequencing of naturally processed peptides bound to HLA molecules provided a means to identify the amino acid residues required for allele-specific epitope-peptide binding (30-32). Data obtained from X-ray crystallographic analysis of HLA-epitope peptide complexes, allowed for the identification and structural characterization of 'binding pockets' within the peptide binding cleft of HLA molecules. More refined epitope anchor motif definitions were then developed using data obtained from *in vitro* peptide-MHC binding assays. It is now well known that the main anchor residues typically occur at position 2 and the carboxyl terminus of peptides 8-11 amino acids in length, thus positions 8, 9, 10 or 11 (33-40). The definition of epitope peptide binding anchor motifs is the key to most, if not all, epitope prediction methods.

[0049] Initial CTL epitope identification methods were developed using common HLA alleles, such as HLA-A2.1. Motifs defined using different HLA molecules were found to be similar and this lead to the definition of HLA supertype families (41). The biological effect of this supertype relationship was first demonstrated for HIV-1 epitopes in a study where the HLA-A3 and -A11 epitope peptide binding patterns repertoires were demonstrated to be overlapping, not only with each other but also with HLA-A31, -A33 and -A*6801 (42). This binding specificity was defined as the HLA-A3 supertype. A significant overlap in peptide binding patterns was also demonstrated amongst several serologically distant HLA-B alleles (43, 44), and multiple HLA-A2 alleles (45, 46),

resulting in the definition of the HLA-B7 and HLA-A2 supertype families. Recognition of epitopes by CTL in a supertype manner has since been demonstrated to occur naturally in infectious diseases and cancer (47-53).

[0050] While only two positions within CTL epitopes are typically characterized as the primary binding anchor positions, the amino acids that can serve as the anchor residues are more variable. The preferred and tolerated amino acids that can serve as anchor residues for the HLA-A2, -A3 and –B7 supertype families of epitopes are listed in Table 1. It is possible for analogous HIV-1 epitope peptides derived from different isolates, which differ with respect to the amino acids used as anchor residues, to bind to HLA molecules similarly. This type of variation can be as conserved since it is likely that CTL produced against one epitope would recognize the related epitope. Thus, variation limited to changes in anchor residues that result in sufficient epitope peptide binding to HLA molecules does not result in immune escape from CTL. Epitopes that contain this type of variation can be identified using the appropriately designed motif search algorithms.

[0051] The TCR of CTL has been reported to be somewhat flexible or promiscuous with respect to recognition of epitope peptides bound to HLA molecules. For HIV-1, this flexibility was demonstrated as CTL recognition of related, but slightly variable, epitopes by single clones of CTL produced following natural infection (54, 55). Similar flexibility of CTL epitope recognition was demonstrated using rhesus macaques and natural infection with SIV or immunization (56, 57). This observation is not unique to HIV-1 and SIV but rather the TCR appears to have evolved to allow promiscuous recognition of peptide epitope bound to MHC molecules (58).

[0052] Selective replacement of certain amino acids in CTL epitope peptides, amino acids thought to represent TCR contact points, is not only tolerated but can increase the recognition of the epitopes by CTL clones (59). The types of amino acid substitutions that can be incorporated, typically amino acids that are similar in chemical properties are best tolerated, and their positions, independent of primary anchor positions, within a selected number of CTL epitopes from tumor associated antigens were also defined.

[0053] For HIV-1 and other infectious agents, reproducible methods for predicting the CTL recognition of related variant epitopes that occur amongst isolates have not been developed. Nor have methods for identifying CTL epitopes that are most likely to induce broadly functional responses when used in vaccine. Thus, there exists a need to develop

such methods to overcome the challenge associated with protein sequence variation in HIV and other infectious agents.

1. Åsjö, B., et al. *AIDS* 11:A17-A36. 1997.

2. Phillips, R. E., et al. *Nature* 354:453-459. 1991.

3. Phillips, R. E., et al. *Chem. Immunol.* 56:150-164. 1993.

4. McMichael, A. J., et al. *Annu. Rev. Immunol.* 15:271-296. 1997.

5. Allen, T. M., et al. *Nature* 407:386-390. 2000.

6. Barouch, D. H., et al. *Nature* 415:335-339. 2002.

7. Evans, D. T., et al. *Nat. Med.* 5:1270-1276. 1999.

8. Vogel, T. U., et al. *J. Virol.* 76:11623-11636. 2002.

9. Nacsa, J., et al. V*irology* 305:210-218. 2003.

10. Borrow, P., et al. *Nat. Med.* 3:205-211. 1997.

11. Price, D. A., et al. *Proc. Natl. Acad.Sci. U.S.A* 94:1890-1895. 1997.

12. Goulder, P. J., et al. *Nat. Med.* 3:212-217. 1997.

13. Goulder, P. J., et al. *Immunol. Lett.* 79:109-116. 2001.

14. Walker, B. D. et al. *Nat. Immunol.* 2:473-475. 2001.

15. Korber, B., et al. *Br. Med. Bull.* 58:19-42. 2001.

16. Gaschen, B., et al. *Science* 296:2354-2360. 2002.

17. Novitsky, V., et al. *J. Virol.* 76:5435-5451. 2002.

18. Davis, N. L., et al. *IUBMB. Life* 53:209-211. 2002.

19. Ellenberger, D. L., et al. *Virology* 302:155-163. 2002.

20. Yusim, K., et al. *J. Virol.* 76:8757-8768. 2002.

21. Ferrari, G., et al. *Proc. Natl. Acad. Sci. U.S.A.* 94:1396-1401. 1997.

22. Fukada, K., et al. *AIDS* 16:701-711. 2002.

23. Ferrari, G., et al. *AIDS Res. Hum. Retroviruses* 16:1433-1443. 2000.

24. Buseyne, F., et al. *Virology* 250:316-324. 1998.

25. Lynch, J. A., et al. *J. Infect. Dis.* 178:1040-1046. 1998.

26. Wilson, S. E., et al. *AIDS Res. Hum. Retroviruses* 14:925-937. 1998.

27. Rowland-Jones, S. L., et al. *Immunol. Lett.* 66:9-14. 1999.

28. Gillespie, G. M., et al. *AIDS* 16:961-972. 2002.

29. Altfeld, M. A., et al. *J. Virol.* 75:1301-1311. 2001.

30. Sette, A. et al. *Curr. Opin. Immunol.* 4:79-86. 1992.

31. Sinigaglia, F. et al. *Curr. Opin. Immunol.* 6:52-56. 1994.

32. Engelhard, V. H. *Curr. Opin. Immunol.* 6:13-23. 1994.

33. Brown, K., et al. *Nature* 364:33-39. 1994.

34. Guo, H. C., et al. *Proc. Natl. Acad. Sci. U.S.A.* 90:8053-8057. 1993.

35. Guo, H. C., et al. *Nature* 360:364-366. 1992.

36. Silver, M. L., et al. *Nature* 360:367-369. 1992.

37. Matsumura, M., et al. *Science* 257:927-934. 1992.

38. Madden, D. R., et al. *Cell* 70:1035-1048. 1995.

39. Fremont, D. H., et al. *Science* 257:919-927. 1992.

40. Sapp, M., et al. *J. Mol. Biol.* 219:277-319. 1991.

41. Kubo, R. T., et al. *J. Immunol.* 152:3913-24. 1994.

42. Sidney, J., et al. *Hum. Immunol.* 45:79-93. 1996.

43. Sidney, J., et al. *J. Immunol.* 154:247-259. 1995.

44. Sidney, J., et al. *J. Immunol.* 157:3480-3490. 1996.

45. Del Guercio, M.-F., et al. *J. Immunol.* 154:685-693. 1995.

46. Fruci, D., et al. *Hum. Immunol.* 38:187-192. 1993.

47. Bertoni, R., et al. *J. Clin. Invest.* 100:503-513. 1997.

48. Threlkeld, S. C., et al. *J. Immunol.* 159:1648-1657.

49. Khanna, R., et al. *J. Virol.* 71:7429-7435. 1997.

50. Bertoletti, A., et al. *Hepatology* 26:1027-1034. 1997.

51. Fleischhauer, K., et al. *J. Immunol.* 157:787-297. 1996.

52. Kawashima, I., et al. *Hum. Immunol.* 59:1-14. 1998.

53. Wang, R. F., et al. *J. Immunol.* 160:890-897. 1998.

54. Tomiyama, H., et al. *J. Immunol.* 30:2521-2530. 2000.

55. Buseyne, F., et al. *Int. Immunol.* 13:941-950. 2001.

56. Shen, L., et al. *J. Immunol.* 153:12-15-1994. 2000.

57. Charini, W. A., et al. *J. Immunol.* 167:4996-5003. 2001.

58. Mason, D. *Immunol. Today.* 19:395-404. 1998.

59. Tangri, S., et al. *J. Exp. Med.* 194:833-846. 2001.

60. Rehermann, B., et al. *J. Exp. Med.* 181:1047-1058. 1995.

61. Chang, K. M., et al. *J. Immunol.* 162:1156-1164. 1999.

62. Chang, K. M., et al. *J. Clin. Invest.* 100:2376-2385. 1997.

63. Propato, A., et al. *Hum. Immunol.* 62:561-576. 2001.

64. Kast, W. M., et al. *J. Immunol.* 152:3904-3912. 1994.

65. Ressing, M. E., et al. *J. Immunol.* 154:5934-5943. 1995.

66. Doolan, D. L., et al. *Immunity.* 7:97-112. 1997.

SUMMARY OF THE INVENTION

[0054] The present invention is directed to methods for selecting a variant of a peptide epitope which induces a CTL response against another variant(s) of the peptide epitope, by determining whether the variant comprises only conserved residues, as defined herein, at non-anchor positions in comparison to the other variant(s).

[0055] In some embodiments, antigen sequences from a population of an infectious agent, said antigens comprising variants of a peptide epitope, are optionally aligned (manually or by computer) along their length, preferably their full length. Variant(s) of a peptide epitope (preferably naturally occurring variants), each 8-11 amino acids in length and comprising the same MHC class I supermotif or motif, are identified manually or with the aid of a computer. In some embodiments, a variant is optionally chosen which comprises preferred anchor residues of said motif and/or which occurs with high frequency within the population of variants. In other embodiments, a variant is randomly chosen. The randomly or otherwise chosen variant is compared to from one to all the remaining variant(s) to determine whether it comprises only conserved residues in the non-anchor positions relative to from one to all the remaining variant(s).

[0056] The present invention is also directed to variants identified by the methods above; peptides comprising such variants; nucleic acids encoding such variants and peptides; cells comprising such variants, and/or peptides, and/or nucleic acids; compositions comprising such variants, and/or peptides, and/or nucleic acids, and/or cells; as well as therapeutic and diagnostic methods for using such variants, peptides, nucleic acids, cells, and compositions.

BRIEF DESCRIPTION OF THE DRAWINGS/FIGURES

[0057] FIGS. 1A-1E. Recognition of variant peptides by CTL generated against a single epitope. Variant peptides were identified from 167 HIV strains for 5 HIV epitopes, 3 HLA-A2 restricted (Env 134, A, Gag 386, B, and Vpr 62, C) and 2 HLA-A11 restricted (Pol 98, D, and Env 47, E). These are listed according to their relationship to a previously determined parent (P) into single anchor substitutions (A), single non-anchor substitutions (NA) or multiple substitutions (M). Binding of each variant peptide is also shown. The number of viral sequences containing each variant peptide is shown in the column labeled # Isolates, and is reported for the total sequences, Clade B sequences (B), and Clade C sequences (C). Finally, the ability of CTL primed against the parent peptide to recognize the variant peptides is shown in the bar graphs.

[0058] FIGS. 2A-2C. Characterization of the peptide-specific T cell lines. A. FACS analysis of the TCRs expressed by peptide -stimulated cells after 0, 1, and 5 peptide stimulations, using a panel of commercially available mAb for mouse TCR 2-14. B-C. Peptide affinity. Parent and variant peptides were titrated against CTL that had been stimulated 5 times with the parent peptide.

[0059] FIGS. 2A-2B. Recognition of a panel of variant peptides by PBL from an HIV-infected individual.

[0060] FIG 4. Prediction of immunological conservation. Gag 271 variants and their binding are shown, along with the number of isolates that express each variant. Immunological recognition was predicted for each variant based on two different choices in the immunizing peptide. On the right, the immunogenicity for each variant is shown.

DETAILED DESCRIPTION OF THE INVENTION

Definitions

[0061] The invention can be better understood with reference to the following definitions:

[0062] An "antigen" refers to a polypeptide encoded by the genome of an infectious agent, or other another source, but preferably an infectious agent in the present invention.

Examples of HIV antigens include Env, Gag, Nef, Pol, Tat, Rev, Vif, Vpr, Vpu, p17, p24, p2, p7, p1, p6, Protease, RT, Integrase, and gp160 (preferably Env, Gag, Nef, Pol, Tat, Rev, Vif, Vpr, Vpu). Examples of HBV antigens include Core, Env, and Pol. Examples of HCV antigens include Core, E1, E2, Ns1, Ns2, Ns3, Ns4, and Ns5. Examples of HPV antigens include E1, E2, E3, E4, E5, E6, E7, L1, and L2. Examples of *Plasmodium falciparum* antigens include CSP, SSP2, Exp1, and LSA1.

[0063] Throughout this disclosure, "binding data" results are often expressed in terms of "IC_{50}'s." IC_{50} is the concentration of peptide in a binding assay at which 50% inhibition of binding of a reference peptide is observed. Given the conditions in which the assays are run (*i.e.*, limiting HLA proteins and labeled peptide concentrations), these values approximate K_D values. Assays for determining binding are described in detail, *e.g.*, in PCT publications WO 94/20127 and WO 94/03205, and other publications such Sidney *et al.*, *Current Protocols in Immunology* 18.3.1 (1998); Sidney, *et al., J. Immunol.* 154:247 (1995); and Sette, *et al., Mol. Immunol.* 31:813 (1994). It should be noted that IC_{50} values can change, often dramatically, if the assay conditions are varied, and depending on the particular reagents used (*e.g.*, HLA preparation, *etc.*). For example, excessive concentrations of HLA molecules will increase the apparent measured IC_{50} of a given ligand.

[0064] Alternatively, binding is expressed relative to a reference peptide. Although as a particular assay becomes more, or less, sensitive, the IC_{50}'s of the peptides tested may change somewhat, the binding relative to the reference peptide will not significantly change. For example, in an assay run under conditions such that the IC_{50} of the reference peptide increases 10-fold, the IC_{50} values of the test peptides will also shift approximately 10-fold. Therefore, to avoid ambiguities, the assessment of whether a peptide is a good (i.e. high), intermediate, weak, or negative binder is generally based on its IC_{50}, relative to the IC_{50} of a standard peptide. The Tables included in this application present binding data in a preferred biologically relevant form of IC_{50} nM.

[0065] Binding may also be determined using other assay systems including those using: live cells (*e.g.*, Ceppellini *et al.*, *Nature* 339:392 (1989); Christnick *et al.*, *Nature* 352:67 (1991); Busch *et al., Int. Immunol.* 2:443 (1990); Hill *et al., J. Immunol.* 147:189 (1991); del Guercio *et al., J. Immunol.* 154:685 (1995)), cell free systems using detergent lysates (*e.g.*, Cerundolo *et al., J. Immunol.* 21:2069 (1991)), immobilized purified MHC (*e.g.*, Hill *et al., J. Immunol.* 152, 2890 (1994); Marshall *et al., J. Immunol.* 152:4946 (1994)), ELISA systems (*e.g.*, Reay *et al., EMBO J.* 11:2829 (1992)), surface plasmon resonance (*e.g.*, Khilko *et al., J. Biol. Chem.* 268:15425 (1993)); high flux soluble phase assays (Hammer *et al., J. Exp. Med.* 180:2353 (1994)), and measurement of class I MHC stabilization or assembly (*e.g.*, Ljunggren *et al., Nature* 346:476 (1990);

Schumacher *et al.*, *Cell* 62:563 (1990); Townsend *et al.*, *Cell* 62:285 (1990); Parker *et al.*, *J. Immunol.* 149:1896 (1992)).

[0066] As used herein, "high affinity" with respect to HLA class I molecules is defined as binding with an IC_{50} or K_D value, of 50 nM or less, "intermediate affinity" is binding with an IC_{50} or K_D value of between 50 and about 500 nM, weak affinity is binding with an IC_{50} or K_D value of between about 500 and about 5000 nM. "High affinity" with repect to binding to HLA class II molecules is defined as binding with an IC_{50} or K_D value of 100 nM or less; "intermediate affinity" is binding with an IC_{50} or K_D value of between about 100 and about 1000 nM.

[0067] A "computer" or "computer system" generally includes: a processor and related computer programs; at least one information storage/retrieval apparatus such as a hard drive, a disk drive or a tape drive; at least one input apparatus such as a keyboard, a mouse, a touch screen, or a microphone; and display structure, such as a screen or a printer. Additionally, the computer may include a communication channel in communication with a network. Such a computer may include more or less than what is listed above.

[0068] "Cross-reactive binding" indicates that a peptide is bound by more than one HLA molecule; a synonym is degenerate binding.

[0069] A "cryptic epitope" elicits a response by immunization with an isolated peptide, but the response is not cross-reactive *in vitro* when intact whole protein, which comprises the epitope, is used as an antigen.

[0070] The term "derived" when used to discuss an epitope is a synonym for "prepared." A derived epitope can be isolated from a natural source, or it can be synthesized in accordance with standard protocols in the art. Synthetic epitopes can comprise artificial amino acids "amino acid mimetics," such as D isomers of natural occurring L amino acids or non-natural amino acids such as cyclohexylalanine. A derived/prepared epitope can be an analog of a native epitope.

[0071] A "diluent" includes sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred diluent for pharmaceutical compositions. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as diluents, particularly for injectable solutions.

[0072] A "dominant epitope" is an epitope that induces an immune response upon immunization with a whole native antigen (see, *e.g.*, Sercarz, *et al.*, *Annu. Rev. Immunol.* 11:729-766, 1993). Such a response is cross-reactive *in vitro* with an isolated peptide epitope.

[0073] An "epitope" is the collective features of a molecule, such as primary, secondary and tertiary peptide structure, and charge, that together form a site recognized by an immunoglobulin, T cell receptor or HLA molecule. Alternatively, an epitope can be defined as a set of amino acid residues which is involved in recognition by a particular immunoglobulin, or in the context of T

cells, those residues necessary for recognition by T cell receptor proteins and/or Major Histocompatibility Complex (MHC) receptors. Epitopes are present in nature, and can be isolated, purified or otherwise prepared/derived by humans. For example, epitopes can be prepared by isolation from a natural source, or they can be synthesized in accordance with standard protocols in the art. Synthetic epitopes can comprise artificial amino acids, "amino acid mimetics," such as D isomers of naturally-occurring L amino acids or non-naturally-occuring amino acids such as cyclohexylalanine. Throughout this disclosure, epitopes may be referred to in some cases as peptides. The variants of the invention are set forth in Tables 6-9 and Figures 1A-4.

[0074] It is to be appreciated that proteins or peptides that comprise a variant of the invention as well as additional amino acid(s) are still within the bounds of the invention. In certain embodiments, the peptide comprises a fragment of an antigen. A "fragment of an antigen" or "antigenic fragment" or simply "fragment" is a portion of an antigen which has 100% identity with a wild type antigen or naturally-ocurring variant thereof. The fragment may or may not comprise an epitope of the invention. The fragment may be less than or equal to 600 amino acids, less than or equal to 500 amino acids, less than or equal to 400 amino acids, less than or equal to 250 amino acids, less than or equal to 100 amino acids, less than or equal to 85 amino acids, less than or equal to 75 amino acids, less than or equal to 65 amino acids, or less than or equal to 50 amino acids in length. In certain embodiments, a fragment is e.g., less than 101 or less than 51 amino acids in length, in any increment down to 5 amino acids in length. For example, the fragment may be 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length.

[0075] In certain embodiments, there is a limitation on the length of a peptide of the invention. The embodiment that is length-limited occurs when the protein/peptide comprising an epitope of the invention comprises a region (i.e., a contiguous series of amino acids) having 100% identity with a native sequence. In order to avoid the definition of epitope from reading, *e.g.*, on whole natural molecules, there is a limitation on the length of any region that has 100% identity with a native peptide sequence. Thus, for a peptide comprising an epitope of the invention and a region with 100% identity with a native peptide sequence, the region with 100% identity to a native sequence generally has a length of: less than or equal to 600 amino acids, often less than or equal to 500 amino acids, often less than or equal to 400 amino acids, often less than or equal to 250 amino acids, often less than or equal to 100 amino acids, often less than or equal to 85 amino acids, often less than or equal to 75 amino acids, often less than or equal to 65 amino acids, and often less than or equal to 50 amino acids. In certain embodiments, an "epitope" of the invention

is comprised by a peptide having a region with less than 51 amino acids that has 100% identity to a native peptide sequence, in any increment down to 5 amino acids.

[0076] Accordingly, peptide or protein sequences longer than 600 amino acids are within the scope of the invention, so long as they do not comprise any contiguous sequence of more than 600 amino acids that have 100% identity with a native peptide sequence. For any peptide that has five contiguous residues or less that correspond to a native sequence, there is no limitation on the maximal length of that peptide in order to fall within the scope of the invention. It is presently preferred that a peptide of the invention (*e.g.*, a peptide comprising an epitope of the invention) be less than 600 residues long in any increment down to eight amino acid residues.

[0077] A peptide epitope occurring with "high frequency" is one that occurs in at least 30%, at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, or at least 90% of the infectious agents in a population. A "high frequency" peptide epitope is one of the more common in a population, preferably the first most common, second most common, third most common, or fourth most common in a population of variant peptide epitopes.

[0078] "Human Leukocyte Antigen" or "HLA" is a human class I or class II Major Histocompatibility Complex (MHC) protein (*see, e.g.*, Stites, *et al.*, IMMUNOLOGY, 8TH ED., Lange Publishing, Los Altos, CA (1994).

[0079] An "HLA supertype or HLA family", as used herein, describes sets of HLA molecules grouped on the basis of shared peptide-binding specificities. HLA class I molecules that share somewhat similar binding affinity for peptides bearing certain amino acid motifs are grouped into such HLA supertypes. The terms HLA superfamily, HLA supertype family, HLA family, and HLA xx-like molecules (where "xx" denotes a particular HLA type), are synonyms. See Tables 1-4.

[0080] As used herein, "high affinity" with respect to HLA class I molecules is defined as binding with an IC_{50}, or K_D value, of 50 nM or less; "intermediate affinity" is binding with an IC_{50} or K_D value of between about 50 and about 500 nM; "weak affinity" is binding with an IC_{50} or K_D value between about 500 and about 5000 nM. "High affinity" with respect to binding to HLA class II molecules is defined as binding with an IC_{50} or K_D value of 100 nM or less; "intermediate affinity" is binding with an IC_{50} or K_D value of between about 100 and about 1000 nM. See "binding data."

[0081] An "IC_{50}" is the concentration of peptide in a binding assay at which 50% inhibition of binding of a reference peptide is observed. Given the conditions in which the assays are run (*i.e.*, limiting HLA proteins and labeled peptide concentrations), these values approximate K_D values. See "binding data."

[0082] The terms "identical" or percent "identity," in the context of two or more peptide sequences or antigen fragments, refer to two or more sequences or subsequences that are the same

or have a specified percentage of amino acid residues that are the same, when compared and aligned for maximum correspondence over a comparison window, as measured using a sequence comparison algorithm or by manual alignment and visual inspection.

[0083] An "immunogenic" peptide or an "immunogenic" epitope or "peptide epitope" is a peptide that comprises an allele-specific motif or supermotif such that the peptide will bind an HLA molecule and induce a CTL and/or HTL response. Thus, immunogenic peptides of the invention are capable of binding to an appropriate HLA molecule and thereafter inducing a cytotoxic T lymphocyte (CTL) response, or a helper T lymphocyte (HTL) response, to the peptide.

[0084] An "infectious agent" refers to a disease-causing microorganism, including viruses, bacteria, fungi, and protozoa against which a cellular immune response, preferably a CTL response, plays a role in acquired immunity. Examples of infectious agents include viruses such as human immunodeficiency virus (HIV), hepatitis B virus (HBV), hepatitis C virus (HCV), human papillomma virus (HPV), Influenza virus, Dengue virus, Epstein-Barr virus, bacteria such as *Mycobacterium tuberculosis* and *Chlamydia,* fungi such as *Candida albicans, Cryptococcus neoformans, Coccidoides spp., Histoplasma spp,* and *Aspergillus fumigatis,* protozoa such as *Plasmodium spp.*, including *P. falciparum, Trypanosoma spp., Schistosoma spp., Leishmania spp* and the like. Preferred infectious agents include HIV, HBV, HCV, HPV, Epstein-Barr virus, *Plasmodium falciparum*, Influenza virus and Dengue virus.

[0085] The phrases "isolated" or "biologically pure" refer to material which is substantially or essentially free from components which normally accompany the material as it is found in its native state. Thus, isolated peptides in accordance with the invention preferably do not contain materials normally associated with the peptides in their *in situ* environment. An "isolated" epitope refers to an epitope that does not include the whole sequence of the antigen or polypeptide from which the epitope was derived. Typically the "isolated" epitope does not have attached thereto additional amino acids that result in a sequence that has 100% identity with a native sequence. The native sequence can be a sequence such as a tumor-associated antigen from which the epitope is derived. Thus, the term "isolated" means that the material is removed from its original environment (*e.g.*, the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or peptide present in a living animal is not isolated, but the same polynucleotide or peptide, separated from some or all of the coexisting materials in the natural system, is isolated. Such a polynucleotide could be part of a vector, and/or such a polynucleotide or peptide could be part of a composition, and still be "isolated" in that such vector or composition is not part of its natural environment. Isolated RNA molecules include *in vivo* or *in vitro* RNA

transcripts of the DNA molecules of the present invention, and further include such molecules produced synthetically.

[0086] "Major Histocompatibility Complex" or "MHC" is a cluster of genes that plays a role in control of the cellular interactions responsible for physiologic immune responses. In humans, the MHC complex is also known as the human leukocyte antigen (HLA) complex. For a detailed description of the MHC and HLA complexes, see, Paul, FUNDAMENTAL IMMUNOLOGY, 3RD ED., Raven Press, New York (1993).

[0087] The term "motif" refers to a pattern of residues in an amino acid sequence of defined length, preferably a peptide of less than about 15 amino acids in length, or less than about 13 amino acids in length, usually from about 8 to about 13 amino acids (*e.g.*, 8, 9, 10, 11, 12, or 13) for a class I HLA motif and from about 6 to about 25 amino acids (*e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25) for a class II HLA motif, which is recognized by a particular HLA molecule. Motifs are typically different for each HLA protein encoded by a given human HLA allele. These motifs often differ in their pattern of the primary and secondary anchor residues. See Tables 1-3.

[0088] A "native" or a "wild type" sequence refers to a sequence found in nature.

[0089] A "negative binding residue" or "deleterious residue" is an amino acid which, if present at certain positions (typically not primary anchor positions) in a peptide epitope, results in decreased binding affinity of the peptide for the peptide's corresponding HLA molecule.

[0090] The term "peptide" is used interchangeably with "oligopeptide" in the present specification to designate a series of residues, typically L-amino acids, connected one to the other, typically by peptide bonds between the α-amino and carboxyl groups of adjacent amino acids.

[0091] A "PanDR binding" peptide or "PADRE®" peptide (Epimmune, San Diego, CA) is a member of a family of molecules that binds more than one HLA class II DR molecule. The pattern that defines the PADRE® family of molecules can be referred to as an HLA Class II supermotif. A PADRE® molecule binds to HLA-DR molecules and stimulates *in vitro* and *in vivo* human helper T lymphocyte (HTL) responses. For a further definition of the PADRE® family, see copending application US serial Nos. 09/709,774, filed November 11, 2000; and 09/707,738, filed November 6, 2000; PCT publication Nos WO 95/07707, and WO 97/26784; U.S. Patent Nos. 5,736,142 issued April 7, 1998; 5,679,640, issued October 21, 1997; and 6,413,935, issued July 2, 2002.

[0092] "Pharmaceutically acceptable" refers to a generally non-toxic, inert, and/or physiologically compatible composition or component of a composition.

[0093] A "pharmaceutical excipient" or "excipient" comprises a material such as an adjuvant, a carrier, pH-adjusting and buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like. A "pharmaceutical excipient" is an excipient which is pharmaceutically acceptable.

[0094] A "primary anchor residue" is an amino acid at a specific position along a peptide sequence which is understood to provide a contact point between the immunogenic peptide and the HLA molecule. One, two or three, primary anchor residues within a peptide of defined length generally defines a "motif" for an immunogenic peptide. These residues are understood to fit in close contact with peptide binding grooves of an HLA molecule, with their side chains buried in specific pockets of the binding grooves themselves. In one embodiment of an HLA class I motif, the primary anchor residues are located at position 2 (from the amino terminal position) and at the carboxyl terminal position of a peptide epitope in accordance with the invention. The primary anchor positions for each motif and supermotif of HLA Class I are set forth in Tables 1-2. For example, analog peptides can be created by altering the presence or absence of particular residues in these anchor positions. Such analogs are used to modulate the binding affinity of an epitope comprising a particular motif or supermotif. A "preferred primary anchor residue" is an anchor residue of a motif or supermotif that is associated with optimal binding. Preferred primary anchor residues are indicated in bold-face in Tables 1-2. A "tolerated primary anchor residue" is an anchor residue of a motif or supermotif that is associated with binding to a lesser extent than a preferred residue. Tolerated primary anchor residues are indicated in italicized text in Tables 1-2.

[0095] "Promiscuous recognition" by a TCR is where a distinct peptide is recognized by the various T cell clones in the context of various HLA molecules. Promiscuous binding by an HLA molecule is synonymous with cross-reactive binding.

[0096] A "protective immune response" or "therapeutic immune response" refers to a CTL and/or an HTL response to an antigen derived from an antigen of an infectious agent, which in some way prevents or at least partially arrests disease symptoms, side effects or progression. The immune response may also include an antibody response which has been facilitated by the stimulation of helper T cells.

[0097] By "ranking" the variants in a population of peptide epitopes is meant ordering each variant by its frequency of occurrance relative to the other variants.

[0098] The term "residue" refers to an amino acid or amino acid mimetic incorporated into a peptide or protein by an amide bond or amide bond mimetic.

[0099] A "secondary anchor residue" is an amino acid at a position other than a primary anchor position in a peptide which may influence peptide binding. A secondary anchor residue occurs at a significantly higher frequency amongst HLA-bound peptides than would be expected by random

distribution of amino acids at a given position. A secondary anchor residue can be identified as a residue which is present at a higher frequency among high or intermediate affinity binding peptides, or a residue otherwise associated with high or intermediate affinity binding. The secondary anchor residues are said to occur at "secondary anchor positions." For example, analog peptides can be created by altering the presence or absence of particular residues in these secondary anchor positions. Such analogs are used to finely modulate the binding affinity of an epitope comprising a particular motif or supermotif. The terminology "fixed peptide" is generally used to refer to an analog peptide that has changes in primary anchore position; not secondary.

[00100] A "subdominant epitope" is an epitope which evokes little or no response upon immunization with a whole antigen or a fragment of the whole antigen comprising a subdominant epitope and a dominant epitope, which comprise the epitope, but for which a response can be obtained by immunization with an isolated peptide, and this response (unlike the case of cryptic epitopes) is detected when whole antigen or a fragment of the whole antigen comprising a subdominant epitope and a dominant epitope is used to recall the response *in vitro* or *in vivo*.

[00101] A "supermotif" is a peptide binding specificity shared by HLA molecules encoded by two or more HLA alleles. Preferably, a supermotif-bearing peptide is recognized with high or intermediate affinity (as defined herein) by two or more HLA antigens.

[00102] "Synthetic peptide" refers to a peptide that is abtained from a non-natural source, *e.g.*, is man-made. Such peptides may be produced using such methods as chemical synthesis or recombinant DNA technology. "Synthetic peptides" include "fusion proteins."

[00103] As used herein, a "vaccine" is a composition used for vaccination, e.g., for prophylaxis or therapy, that comprises one or more peptides of the invention. There are numerous embodiments of vaccines in accordance with the invention, such as by a cocktail of one or more peptides; one or more peptides of the invention comprised by a polyepitopic peptide; or nucleic acids that encode such peptides or polypeptides, *e.g.*, a minigene that encodes a polyepitopic peptide. The "one or more peptides" can include any whole unit integer from 1-150, e.g., at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 or more peptides of the invention. The peptides or polypeptides can optionally be modified, such as by lipidation, addition of targeting or other sequences. HLA class I-binding peptides of the invention can be linked to HLA class II-binding peptides, e.g., a PADRE® universal HTL-bindind peptide, to facilitate activation of both cytotoxic T lymphocytes and helper T lymphocytes. Vaccines can comprise peptide pulsed antigen presenting cells, *e.g.*, dendritic cells.

[00104] A "variant of a peptide epitope" refers to a peptide that is identified from a different viral strain at the same position in an aligned sequence, and that varies by one or

more amino acids from the parent peptide epitope. Examples of peptide epitope variants include those shown in Tables 6-9 and Figures 1A-4. A "variant of an antigen" refers to an antigen that comprises at least one variant of a peptide epitope. Examples of antigen variants include those listed by sequence and/or accession number in Tables 10-22. A "variant of an infectious agent" refers to an infectious agent whose genome encodes at least one variant of an antigen. Variants of infectious agents are related viral, bacterial, funagl, or protozoan strains or isolates that vary in sequence but cause the same disease symptoms. Examples of infectious agent variants include HIV Clade A, B, and C subtypes, HBV subtypes adr, ayr, adw, and ayw, HCV types 1, 2, 3, 4, 5, and 6, HPV strains 1-92 (preferably strains 16, 18, 31, 33, 45, 52, 56, and 58) (see Table 10, listing accession numbers for the complete genome sequences of 167 HIV variants; Table 22, showing an alignment of the complete polyprotein sequences of 50 HCV variants) (see also, *Human Retroviruses and AIDS 2000: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences,* Kuiken CL, et al., Eds. Theoretical Biology and Biophysics Group, Los Alamos National Laboratory, Los Alamos, NM).

[00105] The nomenclature used to describe peptides/proteins follows the conventional practice wherein the amino group is presented to the left (the N-terminus) and the carboxyl group to the right (the C-terminus) of each amino acid residue. When amino acid residue positions are referred to in a peptide epitope they are numbered in an amino to carboxyl direction with position one being the position closest to the amino terminal end of the epitope, or the peptide or protein of which it may be a part. In the formulae representing selected specific embodiments of the present invention, the amino- and carboxyl-terminal groups, although not specifically shown, are in the form they would assume at physiologic pH values, unless otherwise specified. In the amino acid structure formulae, each residue is generally represented by standard three letter or single letter designations. The L-form of an amino acid residue is represented by a capital single letter or a capital first letter of a three-letter symbol, and the D-form for those amino acids having D-forms is represented by a lower case single letter or a lower case three letter symbol. However, when three letter symbols or full names are used without capitals, they may refer to L amino acids. Glycine has no asymmetric carbon atom and is simply referred to as "Gly" or "G". The amino acid sequences of peptides set forth herein are generally designated using the standard single letter symbol. (A, Alanine; C, Cysteine; D, Aspartic Acid; E, Glutamic Acid; F, Phenylalanine; G, Glycine; H, Histidine; I, Isoleucine; K, Lysine; L, Leucine; M, Methionine; N, Asparagine; P, Proline; Q, Glutamine; R, Arginine; S, Serine; T, Threonine; V, Valine; W, Tryptophan; and Y, Tyrosine.) In addition to these symbols, "B"in the single letter abbreviations used herein

designates α-amino butyric acid. In some embodiments, α-amino butyric acid may be replaced with cysteine.

Acronyms used herein are as follows:

APC:	Antigen presenting cell
CD3:	Pan T cell marker
CD4:	Helper T lymphocyte marker
CD8:	Cytotoxic T lymphocyte marker
CEA:	Carcinoembryonic antigen (see, e.g., SEQ ID NO: 363)
CTL:	Cytotoxic T lymphocyte
DC:	Dendritic cells. DC functioned as potent antigen presenting cells by stimulating cytokine release from CTL lines that were specific for a model peptide derived from hepatitis B virus. *In vivo* experiments using DC pulsed *ex vivo* with an HBV peptide epitope have stimulated CTL immune responses *in vivo* following delivery to naïve mice.
DLT:	Dose-limiting toxicity, an adverse event related to therapy.
DMSO:	Dimethylsulfoxide
ELISA:	Enzyme-linked immunosorbant assay
E:T:	Effector:Target ratio
G-CSF:	Granulocyte colony-stimulating factor
GM-CSF:	Granulocyte-macrophage (monocyte)-colony stimulating factor
HBV:	Hepatitis B virus
HER2/neu:	A tumor associated antigen; c-erbB-2 is a synonym (see, e.g., SEQ ID NO: 364)
HLA:	Human leukocyte antigen
HLA-DR:	Human leukocyte antigen class II
HPLC:	High Performance Liquid Chromatography
HTC:	Helper T Cell
HTL:	Helper T Lymphocyte. A synonym for HTC.
ID:	Identity
IFNγ:	Interferon gamma
IL-4:	Interleukin-4
IV:	Intravenous
$LU_{30\%}$:	Cytotoxic activity for 10^6 effector cells required to achieve 30% lysis of a target cell population, at a 100:1 (E:T) ratio.
MAb:	Monoclonal antibody
MAGE:	Melanoma antigen (see, e.g., SEQ ID NO: 365 and 366 for MAGE2 and MAGE3)
MLR:	Mixed lymphocyte reaction
MNC:	Mononuclear cells
PB:	Peripheral blood
PBMC:	Peripheral blood mononuclear cell
ProGP™:	Progenipoietin™ product (Searle, St. Louis, MO), a chimeric flt3/G-CSF receptor agonist.
SC:	Subcutaneous
S.E.M.:	Standard error of the mean
QD:	Once a day dosing
TAA:	Tumor Associated Antigen
TNF:	Tumor necrosis factor
WBC:	White blood cells

[00106] The following describes the peptides, nucleic acid molecules, compositions, and methods of the invention in more detail.

Methods of Identifying Candidate Peptide Epitopes

[00107] The present invention is directed to methods for selecting a variant of a peptide epitope which induces a CTL response against another variant(s) of the peptide epitope, by determining whether the variant comprises only conserved residues, as defined herein, at non-anchor positions in comparison to the other variant(s).

[00108] In some embodiments, antigen sequences from a population of an infectious agent, said antigens comprising variants of a peptide epitope, are optionally aligned (manually or by computer) along their length, preferably their full length. Variant(s) of a peptide epitope (preferably naturally occurring variants), each 8-11 amino acids in length and comprising the same MHC class I supermotif or motif, are identified manually or with the aid of a computer. In some embodiments, a variant is optionally chosen which comprises preferred anchor residues of said motif and/or which occurs with high frequency within the population of variants. In other embodiments, a variant is randomly chosen. The randomly or otherwise chosen variant is compared to from one to all the remaining variant(s) to determine whether it comprises only conserved residues in the non-anchor positions relative to from one to all the remaining variant(s).

[00109] The present invention is also directed to variants identified by the methods above; peptides comprising such variants; nucleic acids encoding such variants and peptides; cells comprising such variants, and/or peptides, and/or nucleic acids; compositions comprising such variants, and/or peptides, and/or nucleic acids, and/or cells; as well as therapeutic and diagnostic methods for using such variants, peptides, nucleic acids, cells, and compositions.

[00110] In some embodiments, the invention is directed to a method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising

a) identifying, from a particular antigen of an infectious agent, variants of a peptide epitope 8-11 amino acids in length, each variant comprising primary anchor residues of the same HLA class I binding motif; and

b) determining whether one of said variants comprises only conserved non-anchor residues in comparison to at least one remaining variant, thereby identifying a candidate peptide epitope.

[00111] In some embodiments, (b) comprises identifying a variant which comprises only conserved non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.

[00112] In some embodiments, the invention is directed to a method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising

a) identifying, from a particular antigen of an infectious agent, variants of a peptide epitope 8-11 amino acids in length, each variant comprising primary anchor residues of the same HLA class I binding motif;

b) determining whether each of said variants comprises conserved, semi-conserved or non-conserved non-anchor residues in comparison to each of the remaining variants; and

c) identifying a variant which comprises only conserved non-anchor residues in comparison to at least one remaining variant.

[00113] In some embodiments, (c) comprises identifying a variant which comprises only conservative non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.

[00114] In some embodiments, the invention is directed to a method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising

a) identifying, from a particular antigen of an infectious agent, a population of variants of a peptide epitope 8-11 amino acids in length, each peptide epitope comprising primary anchor residues of the same HLA class I binding motif;

b) choosing a variant selected from the group consisting of:

i) a variant which comprises preferred primary anchor residues of said motif; and

ii) a variant which occurs with high frequency within the population of variants; and

c) determining whether the variant of (b) comprises only conserved non-anchor residues in comparison to at least one remaining variant, thereby identifying a candidate peptide epitope.

[00115] In some embodiments, (c) comprises identifying a variant which comprises only conservative non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.

[00116] In some embodiments, the invention is directed to method for identifying a candidate peptide epitope which induces a HLA class I CTL response against variants of said peptide epitope, comprising

a) identifying, from a particular antigen of an infectious agent, a population of variants of a peptide epitope 8-11 amino acids in length, each peptide epitope comprising primary anchor residues of the same HLA class I binding motif;

b) choosing a variant selected from the group consisting of:

i) a variant which comprises preferred primary anchor residues of said motif; and

ii) a variant which occurs with high frequency within the population of variants; and

c) determining whether the variant of (b) comprises conserved, semi-conserved or non-conserved non-anchor residues in comparison to each of the remaining variants; and

d) identifying a variant which comprises only conserved non-anchor residues in comparison to at least one remaining variant.

[00117] In some embodiments, (d) comprises identifying a variant which comprises only conservative non-anchor residues in comparison to at least 25%, at least 50%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 97%, or at least 99% of the remaining variants.

[00118] In some embodiments, (a) comprises aligning the sequences of said antigens.

[00119] In some embodiments, (b) comprises comprises choosing a variant which comprises preferred primary anchor residues of said motif.

[00120] In some embodiments, (b) comprises comprises choosing a variant which occurs with high frequency within said population.

[00121] In some embodiments, (b) comprises ranking said variants by frequency of occurrence within said population.

[00122] In some embodiments, (b) comprises choosing a variant which comprises preferred primary anchor residues of said motif and which occurs with high frequency within said population.

[00123] In some embodiments, (b) comprises ranking said variants by frequency of occurrence within said population.

[00124] In some embodiments, the identified variant comprises the fewest conserved anchor residues in comparison to each of the remaining variants.

[00125] In some embodiments, the remaining variants comprise 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 27, 28, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 220, 240, 260, 280, or 300 variants.

[00126] In some embodiments, the infectious agent is selected from the group consisting of: HIV, HBV, HCV, HPV, *Plasmodium falciparum*, Influenza virus, and Dengue virus, Epstein-Barr virus, *Mycobacterium tuberculosis, Chlamydia, Candida albicans, Cryptococcus neoformans, Coccidoides spp., Histoplasma spp, Aspergillus fumigatis, Plasmodium spp., Trypanosoma spp., Schistosoma spp., and Leishmania spp.*

[00127] In some embodiments, the infectious agent is selected from the group consisting of: HIV, HBV, HCV, HPV, *Plasmodium falciparum*, Influenza virus, and Dengue virus.

[00128] In some embodiments, the infectious agent is HIV and the antigen is selected from the group consisting of: Gag, Env, Pol, Nef, Rev, Tat, Vif, Vpr, and Vpu.

[00129] In some embodiments, the infectious agent is HBV and the antigen is selected from the group consisting of: Pol, Env, Core, and NS1/Env2.

[00130] In some embodiments, the infectious agent is HCV and the antigen is selected from the group consisting of: Core, E1, E2, NS1, NS2, NS3, NS4, and NS5.

[00131] In some embodiments, the infectious agent is HPV and the antigen is selected from the group consisting of: E1, E2, E3, E4, E5, E6, E7, L1, and L2.

[00132] In some embodiments, the infectious agent is *Plasmodium falciparum* and the antigen is selected from the group consisting of: CSP, SSP2, EXP1, LSA1.

[00133] In some embodiments, the selected variant and the at least one remaining variant comprise different primary anchor residues of the same motif or supermotif.

[00134] In some embodiments, the motif or supermotif is selected from the group consisting of those in Tables 1-2.

[00135] In some embodiments, the conserved non-anchor residues are at any of positions 3-7 of said variant.

[00136] In some embodiments, the variant comprises only 1-3 conserved non-anchor residues compared to at least one remaining variant.

[00137] In some embodiments, the variant comprises only 1-2 conserved non-anchor residues compared to at least one remaining variant.

[00138] In some embodiments, the variant comprises only 1 conserved non-anchor residue compared to at least one remaining variant.

[00139] In some embodiments, the infectious agent is HPV, and further wherein, the HPV infectious agent is selected from the group consisting of HPV strains 16, 18, 31, 33, 45, 52, 56, and 58.

[00140] In some embodiments, the variants are a population of naturally occurring variants.

[00141] **Optional Alignment.** Optionally, antigen sequences, either full-length or partial, may be aligned mannually or by computer. Convenient computer programs for aligning multiple sequences include Omiga, Oxford software, version 1.1.3, using ClustalW alignment, using an open gap penalty of 10.0, extend gap penalty of 0.05, and delay divergent sequences of 40.0 (See, e.g., Table 21); and BLASTP 2.2.5 (Nov-16-2002) (Altschul, S.F., et al., Nucleic Acids Res. 25:3389-3402 (1997)) using a cutoff = 3e-88 (to select human sequences) (see, e.g., Table 20). Alternatively, alignments may be obtained through publicly available sources such as published journal articles and published patent documents or as disclosed herein (see, e.g., Tables 10-22).

[00142] **HLA Class I Motifs Indicative of CTL Inducing Peptide Epitopes.** A large fraction of HLA class I and class II molecules can be classified into a relatively few supertypes, each respective supertype characterized by largely overlapping peptide binding repertoires, and consensus structures of the main peptide binding pockets. Thus,

peptides of the present invention are preferably identified by the primary residues of any one of several HLA-specific amino acid motifs, or if the presence of the motif corresponds to the ability to bind several allele-specific HLA antigens, a supermotif (*see, e.g.*, Tables 1-2). The preferred primary residues are indicated in bold, while the tolerated primary residues are indicated by italics.

[00143] The primary anchor residues of the HLA class I peptide epitope supermotifs and motifs are summarized in Tables 1-2. Preferred primary anchors are shown in bold, while tolerated primary anchors are shown in italics. Primary and secondary anchor positions for HLA Class I are summarized in Table 3. Allele-specific HLA molecules that fall within the various HLA class I supertypes are listed in Table 4. In some cases, patterns of amino acid residues are present in both a motif and a supermotif. The relationship of a particular motif and any related supermotif is indicated in the description of the individual motifs.

[00144] Thus, the peptide motifs and supermotifs described below, and summarized in Tables 1-2, provide guidance for the identification and use of peptide epitopes comprising primary anchor residues of motifs or supermotifs in accordance with the invention.

[00145] Allele-specific HLA molecules that comprise HLA class I supertype families are listed in Table 4.

[00146] **HLA-A1 supermotif.** The HLA-A1 supermotif is characterized by the presence in peptide ligands of a small (T or S) or hydrophobic (L, I, V, or M) primary anchor residue in position 2, and an aromatic (Y, F, or W) primary anchor residue at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind to the A1 supermotif (*i.e.*, the HLA-A1 supertype) is comprised of at least A*0101, A*2601, A*2602, A*2501, and A*3201 (*see, e.g.*, DiBrino, M. *et al.*, *J. Immunol.* 151:5930, 1993; DiBrino, M. *et al.*, *J. Immunol.* 152:620, 1994; Kondo, A. *et al.*, *Immunogenetics* 45:249, 1997). Other allele-specific HLA molecules predicted to be members of the A1 superfamily are shown in Table 4. Peptides binding to each of the individual HLA proteins can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

[00147] **HLA-A2 supermotif.** Primary anchor specificities for allele-specific HLA-A2.1 molecules (*see, e.g.*, Falk *et al.*, *Nature* 351:290-296, 1991; Hunt *et al.*, *Science* 255:1261-1263, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992; Ruppert *et al.*, *Cell* 74:929-937, 1993) and cross-reactive binding among HLA-A2 and -A28 molecules have been

described. (*See, e.g.*, Fruci *et al.*, *Human Immunol.* 38:187-192, 1993; Tanigaki *et al.*, *Human Immunol.* 39:155-162, 1994; Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994 for reviews of relevant data.) These primary anchor residues define the HLA-A2 supermotif; which presence in peptide ligands corresponds to the ability to bind several different HLA-A2 and -A28 molecules. The HLA-A2 supermotif comprises peptide ligands with L, I, V, M, A, T, or Q as a primary anchor residue at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope.

[00148] The corresponding family of HLA molecules (*i.e.*, the HLA-A2 supertype that binds these peptides) is comprised of at least: A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, and A*6901. Other allele-specific HLA molecules predicted to be members of the A2 superfamily are shown in Table 4. As explained in detail below, binding to each of the individual allele-specific HLA molecules can be modulated by substitutions at the primary anchor and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

[00149] The motifs comprising the primary anchor residues V, A, T, or Q at position 2 and L, I, V, A, or T at the C-terminal position are those most particularly relevant to the invention claimed herein.

[00150] **HLA-A3 supermotif.** The HLA-A3 supermotif is characterized by the presence in peptide ligands of A, L, I, V, M, S, or, T as a primary anchor at position 2, and a positively charged residue, R or K, at the C-terminal position of the epitope, *e.g.*, in position 9 of 9-mers (*see, e.g.*, Sidney *et al.*, *Hum. Immunol.* 45:79, 1996). Exemplary members of the corresponding family of HLA molecules (the HLA-A3 supertype) that bind the A3 supermotif include at least A*0301, A*1101, A*3101, A*3301, and A*6801. Other allele-specific HLA molecules predicted to be members of the A3 supertype are shown in Table 4. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions of amino acids at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

[00151] **HLA-A24 supermotif.** The HLA-A24 supermotif is characterized by the presence in peptide ligands of an aromatic (F, W, or Y) or hydrophobic aliphatic (L, I, V, M, or T) residue as a primary anchor in position 2, and Y, F, W, L, I, or M as primary anchor at the C-terminal position of the epitope (*see, e.g.*, Sette and Sidney, *Immunogenetics*, in press,

1999). The corresponding family of HLA molecules that bind to the A24 supermotif (*i.e.*, the A24 supertype) includes at least A*2402, A*3001, and A*2301. Other allele-specific HLA molecules predicted to be members of the A24 supertype are shown in Table 4. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

[00152] **HLA-B7 supermotif.** The HLA-B7 supermotif is characterized by peptides bearing proline in position 2 as a primary anchor, and a hydrophobic or aliphatic amino acid (L, I, V, M, A, F, W, or Y) as the primary anchor at the C-terminal position of the epitope. The corresponding family of HLA molecules that bind the B7 supermotif (*i.e.*, the HLA-B7 supertype) is comprised of at least twenty six HLA-B proteins including: B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3504, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, and B*7801 (*see, e.g.*, Sidney, *et al., J. Immunol.* 154:247, 1995; Barber, *et al., Curr. Biol.* 5:179, 1995; Hill, *et al., Nature* 360:434, 1992; Rammensee, *et al., Immunogenetics* 41:178, 1995 for reviews of relevant data). Other allele-specific HLA molecules predicted to be members of the B7 supertype are shown in Table 4. As explained in detail below, peptide binding to each of the individual allele-specific HLA proteins can be modulated by substitutions at the primary and/or secondary anchor positions of the peptide, preferably choosing respective residues specified for the supermotif.

[00153] **HLA-B27 supermotif.** The HLA-B27 supermotif is characterized by the presence in peptide ligands of a positively charged (R, H, or K) residue as a primary anchor at position 2, and a hydrophobic (F, Y, L, W, M, I, A, or V) residue as a primary anchor at the C-terminal position of the epitope (*see, e.g.*, Sidney and Sette, *Immunogenetics*, in press, 1999). Exemplary members of the corresponding family of HLA molecules that bind to the B27 supermotif (*i.e.*, the B27 supertype) include at least B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, and B*7301. Other allele-specific HLA molecules predicted to be members of the B27 supertype are shown in Table 4. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

[00154] **HLA-B44 supermotif.** The HLA-B44 supermotif is characterized by the presence in peptide ligands of negatively charged (D or E) residues as a primary anchor in position 2, and hydrophobic residues (F, W, Y, L, I, M, V, or A) as a primary anchor at the C-terminal position of the epitope (*see, e.g.*, Sidney et al., *Immunol. Today* 17:261, 1996). Exemplary members of the corresponding family of HLA molecules that bind to the B44 supermotif (*i.e.*, the B44 supertype) include at least: B*1801, B*1802, B*3701, B*4001, B*4002, B*4006, B*4402, B*4403, and B*4006. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the supermotif.

[00155] **HLA-B58 supermotif.** The HLA-B58 supermotif is characterized by the presence in peptide ligands of a small aliphatic residue (A, S, or T) as a primary anchor residue at position 2, and an aromatic or hydrophobic residue (F, W, Y, L, I, V, M, or A) as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, Sidney and Sette, *Immunogenetics*, in press, 1999 for reviews of relevant data). Exemplary members of the corresponding family of HLA molecules that bind to the B58 supermotif (*i.e.*, the B58 supertype) include at least: B*1516, B*1517, B*5701, B*5702, and B*5801. Other allele-specific HLA molecules predicted to be members of the B58 supertype are shown in Table 4. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

[00156] **HLA-B62 supermotif.** The HLA-B62 supermotif is characterized by the presence in peptide ligands of the polar aliphatic residue Q or a hydrophobic aliphatic residue (L, V, M, I, or P) as a primary anchor in position 2, and a hydrophobic residue (F, W, Y, M, I, V, L, or A) as a primary anchor at the C-terminal position of the epitope (*see, e.g.*, Sidney and Sette, *Immunogenetics*, in press, 1999). Exemplary members of the corresponding family of HLA molecules that bind to the B62 supermotif (*i.e.*, the B62 supertype) include at least: B*1501, B*1502, B*1513, and B5201. Other allele-specific HLA molecules predicted to be members of the B62 supertype are shown in Table 4. Peptide binding to each of the allele-specific HLA molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the supermotif.

[00157] **HLA-A1 motif.** The HLA-A1 motif is characterized by the presence in peptide ligands of T, S, or M as a primary anchor residue at position 2 and the presence of Y as a

primary anchor residue at the C-terminal position of the epitope. An alternative allele-specific A1 motif is characterized by a primary anchor residue at position 3 rather than position 2. This motif is characterized by the presence of D, E, A, or S as a primary anchor residue in position 3, and a Y as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, DiBrino *et al.*, J. Immunol., 152:620, 1994; Kondo *et al.*, *Immunogenetics* 45:249, 1997; and Kubo *et al.*, *J. Immunol.* 152:3913, 1994 for reviews of relevant data). Peptide binding to HLA A1 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

[00158] Those epitopes comprising T, S, or M at position 2 and Y at the C-terminal position are also HLA-A1 supermotif-bearing peptide epitopes, as these residues are a subset of the A1 supermotif primary anchors.

[00159] **HLA-A*0201 motif.** An HLA-A2*0201 motif was determined to be characterized by the presence in peptide ligands of L or M as a primary anchor residue in position 2, and L or V as a primary anchor residue at the C-terminal position of a 9-residue peptide (*see, e.g.*, Falk *et al.*, *Nature* 351:290-296, 1991) and was further found to comprise an I at position 2 and I or A at the C-terminal position of a nine amino acid peptide (*see, e.g.*, Hunt *et al.*, Science 255:1261-1263, March 6, 1992; Parker *et al.*, *J. Immunol.* 149:3580-3587, 1992). The A*0201 allele-specific motif has also been defined by the present inventors to additionally comprise V, A, T, or Q as a primary anchor residue at position 2, and M or T as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, Kast *et al.*, *J. Immunol.* 152:3904-3912, 1994). Thus, the HLA-A*0201 motif comprises peptide ligands with L, I, V, M, A, T, or Q as primary anchor residues at position 2 and L, I, V, M, A, or T as a primary anchor residue at the C-terminal position of the epitope. The preferred and tolerated residues that characterize the primary anchor positions of the HLA-A*0201 motif are identical to the residues describing the A2 supermotif. (For reviews of relevant data, *see, e.g.*, Del Guercio *et al.*, *J. Immunol.* 154:685-693, 1995; Ruppert *et al.*, *Cell* 74:929-937, 1993; Sidney *et al.*, *Immunol. Today* 17:261-266, 1996; Sette and Sidney, *Curr. Opin. in Immunol.* 10:478-482, 1998). Secondary anchor residues that characterize the A*0201 motif have additionally been defined (*see, e.g.*, Ruppert *et al.*, *Cell* 74:929-937, 1993). These are shown in Table 3. Peptide binding to HLA-A*0201 molecules can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

[00160] **HLA-A3 motif.** The HLA-A3 motif is characterized by the presence in peptide ligands of L, M, V, I, S, A, T, F, C, G, or D as a primary anchor residue at position 2, and the presence of K, Y, R, H, F, or A as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, DiBrino *et al.*, *Proc. Natl. Acad. Sci USA* 90:1508, 1993; and Kubo *et al.*, *J. Immunol.* 152:3913-3924, 1994). Peptide binding to HLA-A3 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

[00161] The A3 supermotif primary anchor residues comprise a subset of the A3- and A11-allele specific motif primary anchor residues.

[00162] **HLA-A11 motif.** The HLA-A11 motif is characterized by the presence in peptide ligands of V, T, M, L, I, S, A, G, N, C, D, or F as a primary anchor residue in position 2, and K, R, Y, or H as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, Zhang *et al.*, *Proc. Natl. Acad. Sci USA* 90:2217-2221, 1993; and Kubo *et al.*, *J. Immunol.* 152:3913-3924, 1994). Peptide binding to HLA-A11 can be modulated by substitutions at primary and/or secondary anchor positions, preferably choosing respective residues specified for the motif.

[00163] There is extensive overlap between the A3 and A11 motif primary anchor specificities.

[00164] **HLA-A24 motif.** The HLA-A24 motif is characterized by the presence in peptide ligands of Y, F, W, or M as a primary anchor residue in position 2, and F, L, I, or W as a primary anchor residue at the C-terminal position of the epitope (*see, e.g.*, *Kondo et al.*, *J. Immunol.* 155:4307-4312, 1995; and Kubo *et al.*, *J. Immunol.* 152:3913-3924, 1994). Peptide binding to HLA-A24 molecules can be modulated by substitutions at primary and/or secondary anchor positions; preferably choosing respective residues specified for the motif.

[00165] The primary anchor residues characterizing the A24 allele-specific motif comprise a subset of the A24 supermotif primary anchor residues.

[00166] **Computer or Manual Screening.** Peptides bearing HLA Class I or Class II supermotifs or motifs may be identified by computer searches or manually, e.g., as follows. In utilizing computer screening to identify peptide epitopes, a protein sequence or translated sequence may be analyzed using software developed to search for motifs, for example the "FINDPATTERNS' program (Devereux, *et al. Nucl. Acids Res.* 12:387-395,

1984) or MotifSearch 1.4 software program (D. Brown, San Diego, CA) to identify potential peptide sequences containing appropriate HLA binding motifs. The identified peptides can be scored using customized polynomial algorithms to predict their capacity to bind specific HLA class I or class II alleles. As appreciated by one of ordinary skill in the art, a large array of computer programming software and hardware options are available in the relevant art which can be employed to implement the motifs in order to evaluate (*e.g.*, without limitation, to identify epitopes, identify epitope concentration per peptide length, or to generate analogs) known or unknown peptide sequences.

[00167] Translated antigen protein sequences may be analyzed using a text string search software program, *e.g.*, MotifSearch 1.4 (D. Brown, San Diego) to identify potential peptide sequences containing appropriate HLA binding motifs; alternative programs are readily produced in accordance with information in the art in view of the motif/supermotif disclosure herein. Furthermore, such calculations can be made mentally.

[00168] Identified supermotif or motif sequences may be scored using polynomial algorithms to predict their capacity to bind to specific HLA-Class I or Class II molecules. These polynomial algorithms take into account both extended and refined motifs (that is, to account for the impact of different amino acids at different positions), and are essentially based on the premise that the overall affinity (or ΔG) of peptide-HLA molecule interactions can be approximated as a linear polynomial function of the type:

$$\text{“}\Delta G\text{”} = a_{1i} \times a_{2i} \times a_{3i} \ldots\ldots \times a_{ni}$$

where a_{ji} is a coefficient which represents the effect of the presence of a given amino acid (j) at a given position (i) along the sequence of a peptide of n amino acids. The crucial assumption of this method is that the effects at each position are essentially independent of each other (i.e., independent binding of individual side-chains). When residue j occurs at position i in the peptide, it is assumed to contribute a constant amount j_i to the free energy of binding of the peptide irrespective of the sequence of the rest of the peptide. This assumption is justified by studies from our laboratories that demonstrated that peptides are bound to MHC and recognized by T cells in essentially an extended conformation (data omitted herein).

[00169] The method of derivation of specific algorithm coefficients has been described in Gulukota *et al.*, *J. Mol. Biol.* 267:1258-126, 1997; (see also Sidney *et al.*, *Human Immunol.* 45:79-93, 1996; and Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998). Briefly, for all i positions, anchor and non-anchor alike, the geometric mean of the average

relative binding (ARB) of all peptides carrying j is calculated relative to the remainder of the group, and used as the estimate of j_i. For Class II peptides, if multiple alignments are possible, only the highest scoring alignment is utilized, following an iterative procedure. To calculate an algorithm score of a given peptide in a test set, the ARB values corresponding to the sequence of the peptide are multiplied. If this product exceeds a chosen threshold, the peptide is predicted to bind. Appropriate thresholds are chosen as a function of the degree of stringency of prediction desired.

[00170] Additional methods to identify preferred peptide sequences, which also make use of specific motifs, include the use of neural networks and molecular modeling programs (*see, e.g.*, Milik *et al.*, *Nature Biotechnology* 16:753, 1998; Altuvia *et al.*, *Hum. Immunol.* 58:1, 1997; Altuvia *et al*, *J. Mol. Biol.* 249:244, 1995; Buus, S. *Curr. Opin. Immunol.* 11:209-213, 1999; Brusic, V. *et al.*, *Bioinformatics* 14:121-130, 1998; Parker *et al.*, *J. Immunol.* 152:163, 1993; Meister *et al.*, *Vaccine* 13:581, 1995; Hammer *et al.*, *J. Exp. Med.* 180:2353, 1994; Sturniolo *et al.*, *Nature Biotechnol.* 17:555 1999).

[00171] **Conserved, Semi-conserved, and Non-conserved Non-anchor Residues.** The determination of non-anchor residues as being conserved (conservative) or semi-conserved (semi-conservative) or non-conserved (non-conservative) in comparison to the non-anchor poitions of from one to all of the remaining variant(s) is defined by as follows, the results of which are summarized in Table 5.

[00172] Table 5 shows the similarity assignments between any given amino acid pair so that a given amino acid substitution could be characterized as being a (conservative) or semi-conserved (semi-conservative) or non-conserved (non-conservative) residue.

[00173] The degree of similarity between amino acid pairs was quantified by averaging, for each amino acid pair, the rank coefficient scores for PAM250, hydrophobicity, and side chain volume as described below. Based on the average values of these composite rankings, Table 5 shows each pair to be conserved, semi-conserved or non-conserved.

[00174] The Dayhoff PAM250 score (Dayhoff, M.O., *et al., Atlas of Protein Sequence and Structure*, Vol. 5, suppl.3. (1978) M.O. Dayhoff, ed. National Biomedical Research Foundation, Washington DC, p. 345; Creighton, T.E., *Proteins: structures and molecular properties* (1993) (2nd edition) W.H. Freeman and Company, NY; http://prowl.rockefeller.edu/aainfo/pam250. html) is a commonly utilized protein

alignment scoring matrix which measures the percentage of acceptable point mutations (PAM) within a defined time frame. The frequencies of these mutations are different from what would be expected from the probability of random mutations, and presumably reflect a bias due to the degree of physical and chemical similarity of the amino acid pair involved in the substitution. To obtain a score of amino acid similarity that could be standardized with other measures of similarity, the PAM250 scores were converted to a rank value, where 1 indicates the highest probability of being an accepted mutation.

[00175] The most commonly utilized scales to represent the relative hydrophobicity of the 20 naturally occurring amino acids (Cornette, J., *et al., J. Mol. Biol.* (1987) 195:659) are those developed on the basis of experimental data by Kyte and Doolittle (Kyte, J. and R.F. Doolittle, *J. Mol. Biol.* (1982) 157:105), and by Fauchere and Pliska (Fauchere, J. and V. Pliska, *Eur. J. Med. Chem.* (1983) 18:369). The Kyte/Doolittle scale measures the H_2O/organic solvent partition of individual amino acids. Because it considers the position of amino acids in folded proteins, it may most accurately reflect native hydrophobicity in the context of proteins. The Fauchere/Pliska scale measures the octanol/H_2O partitioning of N-acetyl amino acid amides, and most accurately reflects hydrophobicity in the context of denatured proteins and/or small synthetic peptides. To obtain scores for hydrophobicity, each amino acid residue was ranked on both the Kyte/Doolittle and Fauchere/Pliska hydrophobicity scales. An average rank between the two scales was calculated and the average difference in hydrophobicity for each pair was calculated.

[00176] Finally, for calculating amino acid side-chain volume, the partial volume in solution obtained by noting the increase in volume of water after adding either one molecule or one gram of amino acid residue was considered (Zamyatnin, A.A., *Ann. Rev. Biophys. Bioeng.* (1984) 13:145; Zamyatnin, A.A., *Prog. Biophys. Mol. Biol.* (1972) 24:107). The absolute difference in the partial volume of each possible pairing of the 20 naturally occurring amino acids was calculated and ranked, where 1 indicated residues with the most similar volumes, and 20 the most dissimilar.

[00177] Thus, by consulting Table 5, one can determine whether a residue in a variant is considered to be conserved, semi-conserved, or non-conserved in comparison to a residue in another variant(s). The residue of the parent variant (randomly or otherwise chosen variant) is shown across the top of Table 5, and the residue of the variant(s) it is compared with is shown below the parent residue.

[00178] As shown in Table 5, each of the amino acids shown across the top of the table bears a numerically defined relationship to the remaining 19 genetically encoded amino acids. The lower the index, the higher the conservation; the same amino acid will have a similarity assignment of 1.0; maximally different amino acids will have similarity assignments approaching 20. Using the method set forth above, amino acids which are not gene-encoded can also be assigned similarity indices and can be classified with respect to any natively occurring amino acid as conserved (conservative) or semi-conserved (semi-conservative) or non-conserved (non-conservative).

Variant Peptide Epitopes

[00179] In some embodiments, the invention is directed to an isolated peptide comprising or consisting of a variant. In some embodiments, the invention is directed to an isolated polynucleotide encoding such a peptide.

[00180] The isolated variants of the invention are all class I binding peptides, i.e., CTL peptides. In particular, the variants of the invention comprise a motif or supermotif, as described above. Variants of the invention are those set forth in Tables 6-9 and Figures 1A-4 (SEQ ID Nos:_). Variants of the invention may be referred to herein as "variants" and "variant peptide epitopes" or referred to by Table or referred to by SEQ ID NO. Other peptide epitopes are referred to herein as CTL epitopes or CTL peptides and HTL epitopes or HTL peptides.

[00181] **Peptides and Polynucleotides.** In some embodiments, the invention is directed to an isolated peptide comprising or consisting of a variant, wherein the variant consists of a sequence selected from those in Tables 6-9 and Figures 1A-4 (SEQ ID Nos:__).

[00182] Peptides of the invention may be fusion proteins of variant(s) to CTL epitope(s), and/or HTL epitope(s), and/or linker(s), and/or spacer(s), and/or carrier(s), and/or additional amino acid(s), and/or may comprise or consist of homopolymers of a variant or heteropolymers of more than one variant, as is described in detail below.

[00183] Peptides which comprise a variant of the invention may comprise or consist of a fragment of an antigen ("fragment" or "antigenic fragment"), wherein the fragment comprises a variant. The fragment may be a portion of any antigen of an infectious agent, e.g., the sequences in Tables 11-22 (SEQ ID Nos:__, respectively). The variant of the invention may be within the fragment or may be linked, directly or indirectly, to the fragment.

[00184] The fragment may comprise or consist of a region of a native antigen that contains a high concentration of class I and/or class II epitopes, preferably it contains the greatest number of epitopes per amino acid length. Such epitopes can be present in a frame-shifted manner, *e.g.* a 10

amino acid long peptide could contain two 9 amino acid long epitopes and one 10 amino acid long epitope.

[00185] The fragment may be less than or equal to 600 amino acids, less than or equal to 500 amino acids, less than or equal to 400 amino acids, less than or equal to 250 amino acids, less than or equal to 100 amino acids, less than or equal to 85 amino acids, less than or equal to 75 amino acids, less than or equal to 65 amino acids, or less than or equal to 50 amino acids in length. In certain embodiments, a fragment is less than 101 amino acids in length, in any increment down to 5 amino acids in length. For example, the fragment may be 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 amino acids in length. Fragments of full length antigens may be fragments from about residue 1-20, 21-40, 41-60, 61-80, 81-100, 101-120, 121-140, 141-160, 161-180, 181-200, 201-220, 221-240, 241-260, 261-280, 281-300, 301-320, 321-340, 341-360, 361-380, 381-400, 401-420, 421-440, 441-460, 461-480, 481-500, 501-520, 521-540, 541-560, 561-580, 581-600, 601-620, 621-680, 681-700, 701-720, 721-740, 741-780, 781-800, 801-820, 821-840, 841-860, 861-880, 881-900, 901-920, 921-940, 941-960, 961-980, 981 to the C-terminus of the antigen.

[00186] Peptides which comprise a variant of the invention may be a fusion protein comprising one or more amino acid residues in addition to the variant or fragment. Fusion proteins include homopolymers and heteropolymers, as described below.

[00187] In some embodiments, the peptide comprises or consists of multiple variants, *e.g.*, 2, 3, 4, 5, 6, 7, 8, or 9 variants of the invention. In some embodiments, the peptide comprises at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or at least 8 variants of the invention.

[00188] The peptide may also be a homopolymer of one variant or the peptide may be a heteropolymer which contains at least two different variants. Polymers have the advantage of increased probability for immunological reaction and, where different variants are used to make up the polymer, the ability to induce antibodies and/or T cells that react with different antigenic determinants of the antigen(s) targeted for an immune response.

[00189] A homopolymer may comprise 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 copies of the same variant.

[00190] A heteropolymer may comprise one or more copies of an individual variant and one or more copies of one or more different variants of the invention. The variants that form a heteropolymer may all be from the same antigen, e.g., may be from any of those in Tables 11-22

(SEQ ID NOS:_) or other antigens herein or known in the art, or may be from different antigens, preferably from infectious agents. Combinations of variants that may form a heteropolymer include, for example, Gag 545 variants EPLTSLKSLF (SEQ ID NO:_) and YPLASLKSLF (SEQ ID NO:_), or combinations of peptides from different tables in Tables 6-9 and/or Figures 1A-4 or those combinations in Tables 23-28. Heteropolymers may contain multiple copies of one or more variants.

[00191] Thus, peptides of the invention such as heteropolymers may comprise a first variant and at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 other (different) variants.

[00192] In some embodiments, the peptide comprising a variant may also comprise a number of CTL and/or HTL epitopes, *e.g.*, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 CTL and/or HTL epitopes.

[00193] The CTL and/or HTL epitope and the variant of the invention may be from the same antigen of an infectious agent or from different antigens. Thus, for example, if the variant is from HIV pol, the CTL peptide and/or HTL peptide may also be from HIV pol. Alternatively, if the variant is from HIV pol, the CTL peptide and/or HTL peptide may be from another antigen such as HIV env or HIV vpr. As another example, if the variant is from HBV E6, the CTL peptide and/or HTL peptide may be from HBV E7. The CTL and/or HTL epitope and the variant of the invention may be from the same infectious agent or different infectious agents. Thus, for example, the variant may be from HIV, and the CTL and/or HTL epitope may be from HIV or may be from another infectious agent sush such as HBV, HCV, HPV, or *Plasmodium falciparum*.

[00194] The CTL peptide and/or HTL peptide may be from other antigens including hepatitis B core and surface antigens (HBVc, HBVs), hepatitis C antigens, Epstein-Barr virus antigens, human immunodeficiency virus (HIV) antigens and human papilloma virus (HPV) antigens (in particular anitgens from HPV-16, HPV-18, HPV-31, HPV-33, HPV-45, HPV-52, HPV-56 and HPV-58, *Mycobacterium tuberculosis* and *Chlamydia*. Examples of suitable fungal antigens include those derived from *Candida albicans, Cryptococcus neoformans, Coccidoides spp., Histoplasma spp,* and *Aspergillus fumigatis*. Examples of suitable protozoan parasitic antigens include those derived from *Plasmodium spp.*, including *P. falciparum, Trypanosoma spp., Schistosoma spp., Leishmania spp* and the like.

[00195] Alternatively, the CTL peptide and/or HTL peptide may be from tumor-associated antigens such as but not limited to, melanoma antigens MAGE-1, MAGE-2, MAGE-3, MAGE-11, MAGE-A10, as well as BAGE, GAGE, RAGE, MAGE-C1, LAGE-1, CAG-3, DAM, MUC1, MUC2, MUC18, NY-ESO-1, MUM-1, CDK4, BRCA2, NY-LU-1, NY-LU-7, NY-LU-12,

CASP8, RAS, KIAA-2-5, SCCs, p53, p73, CEA, HER2/neu, Melan-A, gp100, tyrosinase, TRP2, gp75/TRP1, kallikrein, prostate-specific membrane antigen (PSM), prostatic acid phosphatase (PAP), prostate-specific antigen (PSA), PT1-1, ∃-catenin, PRAME, Telomerase, FAK, cyclin D1 protein, NOEY2, EGF-R, SART-1, CAPB, HPVE7, p15, Folate receptor CDC27, PAGE-1, and PAGE-4.

[00196] Examples of CTL peptides and HTL peptides are disclosed in WO 01/42270, published 14 June 2001; WO 01/41788, published 14 June 2001; WO 01/42270, published 14 June 2001; WO 01/45728, published 28 June 2001; and WO 01/41787, published 14 June 2001.

[00197] The HTL peptide may comprise a "loosely HLA-restricted" or "promiscuous" sequence. Examples of amino acid sequences that are promiscuous include sequences from antigens such as tetanus toxoid at positions 830-843 (QYIKANSKFIGITE; SEQ ID NO: 627), *Plasmodium falciparum* CS protein at positions 378-398 (DIEKKIAKMEKASSVFNVVNS; SEQ ID NO: 628), and Streptococcus 18kD protein at positions 116-131 (GAVDSILGGVATYGAA; SEQ ID NO: 629). Other examples include peptides bearing a DR 1-4-7 supermotif, or either of the DR3 motifs.

[00198] The HTL peptide may comprise a synthetic peptide such as a Pan-DR-binding epitope (*e.g.*, a PADRE® peptide, Epimmune Inc., San Diego, CA, described, for example, in U.S. Patent Number 5,736,142), for example, having the formula aKXVAAZTLKAAa, where "X" is either cyclohexylalanine, phenylalanine, or tyrosine; "Z" is either tryptophan, tyrosine, histidine or asparagine; and "a" is either D-alanine or L-alanine (SEQ ID NO: 746). Certain pan-DR binding epitopes comprise all "L" natural amino acids; these molecules can be provided as peptides or in the form of nucleic acids that encode the peptide. See also, U.S. Patent Nos. 5,679,640 and 6,413,935.

[00199] The peptide comprising a variant may comprise additional amino acid(s). Such additional amino acids may be Ala, Arg, Asn, Asp, Cys, Gln, Gly, Glu, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Tyr, Trp, Val, amino acid mimetics, and other unnatural amino acids such as those described below. Additional amino acids may provide for ease of linking peptides one to another, for linking variants to one another, for linking variants to CTL and/or HTL epitopes, for coupling to a carrier support or larger peptide, for modifying the physical or chemical properties of the peptide or oligopeptide, or the like. Amino acids such as Ala, Arg, Asn, Asp, Cys, Gln, Gly, Glu, His, Ile, Leu, Lys, Met, Phe,

Pro, Ser, Thr, Tyr, Trp, or Val, or the like, can be introduced at the C- and/or N-terminus of the peptide and/or can be introduced internally.

[00200] The peptide comprising a variant may comprise an amino acid spacer(s), which may be joined to the variants, CTL epitopes, HTL epitopes, carriers, etc. within a peptide or may be joined to the peptide at the N-and/or C-terminus. Thus, spacers may be at the N-terminus or C-terminus of peptide, or may be internal such that they link or join variants, CTL epitopes, HTL epitopes, carriers, additional amino acids, and/or antigenic fragments one to the other.

[00201] The spacer is typically comprised of one or more relatively small, neutral molecules, such as amino acids or amino acid mimetics, which are substantially uncharged under physiological conditions. The spacers are typically selected from, *e.g.*, Ala, Gly, or other neutral spacers of nonpolar amino acids or neutral polar amino acids. It will be understood that the optionally present spacer may be composed of the same residues or may be composed of one or more different residues and thus may be a homo- or hetero-oligomer of spacer residues. Thus, the spacer may contain more than one Ala residue (poly-alanine) or more than one Gly residue (poly-glycine), or may contain both Ala and Gly residues, e.g., Gly, Gly-Gly-, Ser,Ser-Ser-, Gly-Ser-, Ser-Gly-, etc. When present, the spacer will usually be at least one or two residues, more usually three to six residues and sometimes 10 or more residues, e.g., 3, 4, 5, 6, 7, 8, 9, or 10, or even more residues. (Livingston, B.D. *et al. Vaccine* 19:4652-4660 (2000)).

[00202] Peptides comprising a variant may comprise carrier(s) such as those well known in the art, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza virus proteins, hepatitis B virus core protein, and the like. (See Table 29).

[00203] In addition, the peptide comprising or consisting of a variant may be modified by terminal-NH_2 acylation, *e.g.*, by alkanoyl (C_1-C_{20}) or thioglycolyl acetylation, terminal-carboxyl amidation, *e.g.*, ammonia, methylamine, *etc.* In some instances these modifications may provide sites for linking to a support or other molecule.

[00204] The peptides in accordance with the invention can contain modifications such as but not limited to glycosylation, side chain oxidation, biotinylation, phosphorylation, addition of a surface active material, *e.g.* a lipid, or can be chemically modified, *e.g.*,

acetylation, *etc.* Moreover, bonds in the peptide can be other than peptide bonds, *e.g.*, covalent bonds, ester or ether bonds, disulfide bonds, hydrogen bonds, ionic bonds, *etc.*

[00205] Peptides of the present invention may contain substitutions to modify a physical property (*e.g.*, stability or solubility) of the resulting peptide. For example, peptides may be modified by the substitution of a cysteine (C) with α-amino butyric acid ("B"). Due to its chemical nature, cysteine has the propensity to form disulfide bridges and sufficiently alter the peptide structurally so as to reduce binding capacity. Substituting α-amino butyric acid for C not only alleviates this problem, but actually improves binding and crossbinding capability in certain instances. Substitution of cysteine with α-amino butyric acid may occur at any residue of a peptide, *e.g.*, at either anchor or non-anchor positions of a variant within a peptide, or at other positions of a peptide.

[00206] The peptides comprising a variant can comprise amino acid mimetics or unnatural amino acids, *e.g.* D- or L-naphylalanine; D- or L-phenylglycine; D- or L-2-thieneylalanine; D- or L-1, -2, 3, or 4-pyreneylalanine; D- or L-3 thieneylalanine; D- or L-(2-pyridinyl)-alanine; D- or L-(3-pyridinyl)-alanine; D- or L-(2-pyrazinyl)-alanine; D- or L-(4-isopropyl)-phenylglycine; D-(trifluoromethyl)-phenylglycine; D-(trifluoromethyl)-phenylalanine; D-ρ-fluorophenylalanine; D- or L-ρ-biphenylphenylalanine; D- or L-ρ-methoxybiphenylphenylalanine; D- or L-2-indole(alkyl)alanines; and, D- or L-alkylalanines, where the alkyl group can be a substituted or unsubstituted methyl, ethyl, propyl, hexyl, butyl, pentyl, isopropyl, iso-butyl, sec-isotyl, iso-pentyl, or a non-acidic amino acids. Aromatic rings of a non-natural amino acid include, *e.g.*, thiazolyl, thiophenyl, pyrazolyl, benzimidazolyl, naphthyl, furanyl, pyrrolyl, and pyridyl aromatic rings. Modified peptides that have various amino acid mimetics or unnatural amino acids are particularly useful, as they tend to manifest increased stability *in vivo*. Such peptides may also possess improved shelf-life or manufacturing properties.

[00207] Peptide stability can be assayed in a number of ways. For instance, peptidases and various biological media, such as human plasma and serum, have been used to test stability. *See, e.g.*, Verhoef, *et al.*, *Eur. J. Drug Metab. Pharmacokinetics* 11:291 (1986). Half-life of the peptides of the present invention is conveniently determined using a 25% human serum (v/v) assay. The protocol is generally as follows: Pooled human serum (Type AB, non-heat inactivated) is delipidated by centrifugation before use. The serum is then diluted to 25% with RPMI-1640 or another suitable tissue culture medium. At predetermined time intervals, a small amount of reaction solution is removed and added to

either 6% aqueous trichloroacetic acid (TCA) or ethanol. The cloudy reaction sample is cooled (4°C) for 15 minutes and then spun to pellet the precipitated serum proteins. The presence of the peptides is then determined by reversed-phase HPLC using stability-specific chromatography conditions.

[00208] As indicated above, the peptides in accordance with the invention can be a variety of lengths, and either in their neutral (uncharged) forms or in forms which are salts. The peptides in accordance with the invention can contain modifications such as glycosylation, side chain oxidation, or phosphorylation, generally subject to the condition that modifications do not destroy the biological activity of the peptides.

[00209] The peptides of the invention may be lyophylized, or may be in crystal form.

[00210] It is generally preferable that the variant peptide epitope be as small as possible while still maintaining substantially all of the immunologic activity of the native protein. When possible, it may be desirable to optimize HLA class I binding epitopes of the invention to a length of about 8 to about 13 amino acid residues, for example, 8, 9, 10, 11, 12 or 13, preferably 8 to 11 or 9 to 10. It is to be appreciated that one or more epitopes in this size range can be comprised by a longer peptide (see the Definition Section for the term "epitope" for further discussion of peptide length). HLA class II binding epitopes are preferably optimized to a length of about 6 to about 30 amino acids in length, *e.g.*, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 or 30, preferably to between about 13 and about 20 residues, *e.g.*, 13, 14, 15, 16, 17, 18, 19 or 20. Preferably, the epitopes are commensurate in size with endogenously processed pathogen-derived peptides or tumor cell peptides that are bound to the relevant HLA molecules. The identification and preparation of peptides of various lengths can be carried out using the techniques described herein.

[00211] Peptides in accordance with the invention can be prepared synthetically, by recombinant DNA technology or chemical synthesis, or can be isolated from natural sources such as native tumors or pathogenic organisms. Epitopes may be synthesized individually or joined directly or indirectly in a peptide. Although the peptide will preferably be substantially free of other naturally occurring host cell proteins and fragments thereof, in some embodiments the peptides may be synthetically conjugated to be joined to native fragments or particles.

[00212] The peptides of the invention can be prepared in a wide variety of ways. For relatively short sizes, the peptides can be synthesized in solution or on a solid support in

accordance with conventional techniques. Various automatic synthesizers are commercially available and can be used in accordance with known protocols. (*See*, for example, Stewart & Young, SOLID PHASE PEPTIDE SYNTHESIS, 2D. ED., Pierce Chemical Co., 1984). Further, individual peptides can be joined using chemical ligation to produce larger peptides that are still within the bounds of the invention.

[00213] Alternatively, recombinant DNA technology can be employed wherein a nucleotide sequence which encodes a peptide inserted into an expression vector, transformed or transfected into an appropriate host cell and cultivated under conditions suitable for expression. These procedures are generally known in the art, as described generally in Sambrook *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989). Thus, recombinant peptides, which comprise or consist of one or more epitopes of the invention, can be used to present the appropriate T cell epitope.

[00214] Polynucleotides encoding each of the peptides above are also part of the invention. As appreciated by one of ordinary skill in the art, various nucleic acids will encode the same peptide due to the redundancy of the genetic code. Each of these nucleic acids falls within the scope of the present invention. This embodiment of the invention comprises DNA and RNA, and in certain embodiments a combination of DNA and RNA. It is to be appreciated that any polynucleotide that encodes a peptide in accordance with the invention falls within the scope of this invention.

[00215] The polynucleotides encoding peptides contemplated herein can be synthesized by chemical techniques, for example, the phosphotriester method of Matteucci, *et al., J. Am. Chem. Soc.* 103:3185 (1981). Polynucleotides encoding peptides comprising or consisting of a variant can be made simply by substituting the appropriate and desired nucleic acid base(s) for those that encode a related (e.g., analogous) epitope.

[00216] The polynucleotide, e.g. minigene (see below), may be produced by assembling oligonucleotides that encode the plus and minus strands of the polynucleotide, e.g. minigene. Overlapping oligonucleotides (15-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. A polynucleotide, e.g. minigene, encoding the peptide of the invention, can be cloned into a desired vector such as an expression vector. The coding sequence can then be provided with appropriate linkers and ligated into expression vectors commonly

available in the art, and the vectors used to transform suitable hosts to produce the desired peptide such as a fusion protein.

[00217] A large number of such vectors and suitable host systems are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example. Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, phagescript, psiX174, pBluescript SK, pbsks, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); pCR (Invitrogen). Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia); p75.6 (valentis); pCEP (Invitrogen); pCEI (Epimmune). However, any other plasmid or vector can be used as long as it is replicable and viable in the host.

[00218] As representative examples of appropriate hosts, there can be mentioned: bacterial cells, such as *E. coli, Bacillus subtilis, Salmonella typhimurium* and various species within the genera Pseudomonas, Streptomyces, and Staphylococcus; fungal cells, such as yeast; insect cells such as Drosophila and Sf9; animal cells such as COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell 23*:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines or Bowes melanoma; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

[00219] Thus, the present invention is also directed to vectors, preferably expression vectors useful for the production of the peptides of the present invention, and to host cells comprising such vectors.

[00220] Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which can be, for example, a cloning vector or an expression vector. The vector can be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the polynucletides. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

[00221] For expression of the peptides, the coding sequence will be provided with operably linked start and stop codons, promoter and terminator regions and usually a replication system to provide an expression vector for expression in the desired cellular host. For

example, promoter sequences compatible with bacterial hosts are provided in plasmids containing convenient restriction sites for insertion of the desired coding sequence. The resulting expression vectors are transformed into suitable bacterial hosts.

[00222] Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, *e.g.*, the ampicillin resistance gene of *E. coli* and *S. cerevisiae* TRP1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK), ∀-factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, *e.g.*, stabilization or simplified purification of expressed recombinant product.

[00223] Yeast, insect or mammalian cell hosts may also be used, employing suitable vectors and control sequences. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by Gluzman, *Cell 23*:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. Such promoters may also be derived from viral sources, such as, e.g., human cytomegalovirus (CMV-IE promoter) or herpes simplex virus type-1 (HSV TK promoter). Nucleic acid sequences derived from the SV40 splice, and polyadenylation sites can be used to provide the required nontranscribed genetic elements.

[00224] Polynucleotides encoding peptides of the invention may also comprise a ubiquitination signal sequence, and/or a targeting sequence such as an endoplasmic reticulum (ER) signal sequence to facilitate movement of the resulting peptide into the endoplasmic reticulum.

[00225] Polynucleotides of the invention, e.g., minigenes, may be expressed in human cells. A human codon usage table can be used to guide the codon choice for each amino

acid. Such polynucleotides preferably comprise spacer amino acid residues between variants, such as those described above, or may comprise naturally-occurring flanking sequences adjacent to the variants (and/or CTL and HTL epitopes).

[00226] The peptides of the invention can also be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. As an example of this approach, vaccinia virus is used as a vector to express nucleotide sequences that encode the peptides of the invention. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the polypeptides of the invention, *e.g.* adeno and adeno-associated virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, will be apparent to those skilled in the art from the description herein. A preferred vector is Modified Vaccinia Ankara (MVA) (*e.g.*, Bavarian Noridic (MVA-BN)).

[00227] Standard regulatory sequences well known to those of skill in the art are preferably included in the vector to ensure expression in the human target cells. Several vector elements are desirable: a promoter with a downstream cloning site for polynucleotide, *e.g.*, minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (*e.g.* ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, *e.g.*, the human cytomegalovirus (hCMV) promoter. See, *e.g.*, U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences. A preferred promoter is the CMV-IE promoter.

[00228] Polynucleotides, e.g. minigenes, may comprise one or more synthetic or naturally-occurring introns in the transcribed region. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing polynucleotide, e.g. minigene, expression.

[00229] In addition, the polynucleotide, e.g. minigene, may comprise immunostimulatory sequences (ISSs or CpGs). These sequences may be included in the vector, outside the polynucleotide (e.g. minigene) coding sequence to enhance immunogenicity.

[00230] In some embodiments, a bi-cistronic expression vector which allows production of both the polynucleotide- (e.g. minigene-) encoded peptides of the invention and a second protein (*e.g.*, one that modulates immunogenicity) can be used. Examples of proteins or

polypeptides that, if co-expressed with peptides of the invention, can enhance an immune response include cytokines (*e.g.*, IL-2, IL-12, GM-CSF), cytokine-inducing molecules (*e.g.*, LeIF), costimulatory molecules, or pan-DR binding proteins (PADRE® molecules, Epimmune, San Diego, CA). Helper T cell (HTL) epitopes such as PADRE® molecules can be joined to intracellular targeting signals and expressed separately from expressed peptides of the invention. Specifically decreasing the immune response by co-expression of immunosuppressive molecules (*e.g.* TGF-β) may be beneficial in certain diseases.

[00231] Once an expression vector is selected, the polynucleotide, e.g. minigene, is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate bacterial strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the polynucleotide, e.g. minigene, as well as all other elements included in the vector, are confirmed using restriction mapping, DNA sequence analysis, and/or PCR analysis. Bacterial cells harboring the correct plasmid can be stored as cell banks.

[00232] Therapeutic/prophylactic quantities of DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate growth medium, and are grown to saturation in shaker flasks or a bioreactor according to well known techniques. Plasmid DNA is purified using standard bioseparation technologies such as solid phase anion-exchange resins available, *e.g.*, from QIAGEN, Inc. (Valencia, California). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

[00233] Purified polynucleotides, e.g. minigenes, can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized polynucleotide, e.g. DNA, in sterile phosphate-buffer saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of polynucleotide vaccines, alternative methods of formulating purified plasmid DNA may be used. A variety of such methods have been described, and new techniques may become available. Cationic lipids, glycolipids, and fusogenic liposomes can also be used in the formulation (see, *e.g.*, WO 93/24640; Mannino & Gould-Fogerite, *BioTechniques* 6(7): 682 (1988); U.S. Patent No. 5,279,833; WO 91/06309; and Felgner, *et al., Proc. Nat'l Acad. Sci. USA* 84:7413 (1987). In addition, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) can also be complexed to purified plasmid DNA to

influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

[00234] Known methods in the art can be used to enhance delivery and uptake of a polynucleotide *in vivo*. For example, the polynucleotide can be complexed to polyvinylpyrrolidone (PVP), to prolong the localized bioavailability of the polynucleotide, thereby enhancing uptake of the polynucleotide by the organisum (*see e.g.*, U.S. Patent No. 6,040,295; EP 0 465 529; WO 98/17814). PVP is a polyamide that is known to form complexes with a wide variety of substances, and is chemically and physiologically inert.

[00235] Target cell sensitization can be used as a functional assay of the expression and HLA class I presentation of polynucleotide- (e.g. minigene-) encoded peptides. For example, the polynucleotide, e.g. plasmid DNA, is introduced into a mammalian cell line that is a suitable target for standard CTL chromium release assays. The transfection method used will be dependent on the final formulation. For example, electroporation can be used for "naked" DNA, whereas cationic lipids or PVP-formulated DNA allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). The transfected cells are then chromium-51 (^{51}Cr) labeled and used as targets for epitope-specific CTLs. Cytolysis of the target cells, detected by ^{51}Cr release, indicates both production and HLA presentation of, polynucleotide-, e.g. minigene-, encoded variants of the invention, or peptides comprising them. Expression of HTL epitopes may be evaluated in an analogous manner using assays to assess HTL activity.

[00236] *In vivo* immunogenicity is a second approach for functional testing of polynucleotides, e.g. minigenes. Transgenic mice expressing appropriate human HLA proteins are immunized with the polynucleotide, e.g. DNA, product. The dose and route of administration are formulation dependent (*e.g.*, IM for polynucleotide (*e.g.*, naked DNA or PVP-formulated DNA) in PBS, intraperitoneal (IP) for lipid-complexed polynucleotide (*e.g.*, DNA)). Eleven to twenty-one days after immunization, splenocytes are harvested and restimulated for one week in the presence of polynucleotides encoding each peptide being tested. Thereafter, for peptides comprising or consisting of variants, standard assays are conducted to determine if there is cytolysis of peptide-loaded, ^{51}Cr-labeled target cells. Once again, lysis of target cells that were exposed to variants corresponding to those encoded by the polynucleotide (e.g. minigene) demonstrates polynucleotide (e.g., DNA)

vaccine function and induction of CTLs. Immunogenicity of HTL epitopes is evaluated in transgenic mice in an analogous manner.

[00237] Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Using this technique, particles comprised solely of a polynucleotide such as DNA are administered. In a further alternative embodiment for ballistic delivery, polynucleotides such as DNA can be adhered to particles, such as gold particles.

[00238] The use of polynucleotides such as multi-epitope minigenes is described herein and in, *e.g.* co-pending application U.S.S.N. 09/311,784; Ishioka *et al.*, *J. Immunol.* 162:3915-3925, 1999; An, L. and Whitton, J. L., *J. Virol.* 71:2292, 1997; Thomson, S. A. *et al.*, *J. Immunol.* 157:822, 1996; Whitton, J. L. *et al.*, *J. Virol.* 67:348, 1993; Hanke, R. *et al.*, *Vaccine* 16:426, 1998. For example, a polynucleotide such as a multi-epitope DNA plasmid can be engineered which encodes an epitope derived from multiple regions of a infectious agent (*e.g.*, p53, HER2/nev, MAGE-2/3, or CEA), a pan-DR binding peptide such as the PADRE® universal helper T cell epitope, and an endoplasmic reticulum-translocating signal sequence. As descibed in the sections above, a peptide/polynucleotide may also comprise/encode epitopes that are derived from other infectious agents.

[00239] Thus, the invention includes peptides as described herein, polynucleotides encoding each of said peptides, as well as compositions comprising the peptides and polynucleotides, and includes methods for producing and methods of using the peptides, polynucleotides, and compositions, as further described below.

[00240] **Compositions.** In other embodiments, the invention is directed to a composition comprising one or more peptides and/or polynucleotides of the invention and optionally another component(s).

[00241] In some embodiments, the composition comprises or consists of multiple peptides, *e.g.*, 2, 3, 4, 5, 6, 7, 8, or 9 peptides of the invention. In some embodiments, the composition comprises at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or at least 8 peptides of the invention. Combinations of peptides include, for example, a peptide comprising or alternatively consisting of the Gag 545 variant EPLTSLKSLF (SEQ ID NO:_) and a peptide comprising or alternatively consisting of the Gag 545 variant YPLASLKSLF (SEQ ID NO:_), or combinations of peptides from different tables in Tables 6-9 and/or Figures 1A-4.

[00242] Compositions of the invention may comprise polynucleotides encoding the above peptides and/or combinations of peptides.

[00243] The composition can comprise at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, or at least 8 peptides and/or polynucleotides selected from those described above or below. At least one of the one or more peptides can be a heteropolymer or a homopolymer. Additionally, the composition can comprise a CTL and/or HTL epitope, which can be derived from a tumor-associated antigen. The additional epitope can also be a PanDR binding molecule, (*e.g.*, a PADRE® universal helper T cell epitope).

[00244] Optional components include excipients, diluents, proteins such as peptides comprising a CTL epitope, and/or an HTL epitope such as a pan-DR binding peptide (*e.g.*, a PADRE® universal helper T cell epitope), and/or a carrier, polynucleotides encoding such proteins, lipids, or liposomes, as well as other components described herein. There are numerous embodiments of compositions in accordance with the invention, such as a cocktail of one or more peptides and/or polynucleotides (*e.g.*, minigenes); a cocktail of one or more peptides and/or polynucleotides (*e.g.*, minigenes) and one or more CTL and/or HTL epitopes.

[00245] Compositions may comprise one or more peptides (and/or polynucleotides such as minigenes) of the invention, along with one or more other components as described above and herein. "One or more" refers to any whole unit integer from 1-150, *e.g.*, at least 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 , 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, or 150 peptides, polynucleotides, or other components.

[00246] Compositions of the invention may be, for example, polynucleotides or polypeptides of the invention combined with or complexed to cationic lipid formulations; lipopeptides (*e.g.*,Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), encapsulated *e.g.*, in poly(DL-lactide-co-glycolide) ("PLG") microspheres (*see, e.g.*, Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, 1991: Alonso *et al.*, *Vaccine* 12:299-306, 1994; Jones *et al.*, *Vaccine* 13:675-681, 1995); peptide compositions contained in immune stimulating complexes (ISCOMS) (*see, e.g.*, Takahashi *et al.*, *Nature* 344:873-875, 1990; Hu *et al.*, *Clin Exp Immunol.* 113:235-243, 1998); multiple antigen peptide systems (MAPs) (*see e.g.*, Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996); viral, bacterial, or, fungal delivery vectors (Perkus, M. E. *et*

al., In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F. H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990); particles of viral or synthetic origin (*e.g.*, Kofler, N. *et al.*, *J. Immunol. Methods.* 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Falo, L. D., Jr. *et al.*, *Nature Med.* 7:649, 1995); adjuvants (*e.g.*, incomplete Freund's adjuvant) (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993); liposomes (Reddy, R. *et al.*, *J. Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996); or, particle-absorbed cDNA or other polynucleotides of the invention (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993), etc. Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) or attached to a stress protein, e.g., HSP 96 (Stressgen Biotechnologies Corp., Victoria, BC, Canada) can also be used.

[00247] Compositions of the invention comprise polynucleotide-mediated modalities. DNA or RNA encoding one or more of the peptides of the invention can be administered to a patient. This approach is described, for instance, in Wolff *et. al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; and, WO 98/04720. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivicaine, polymers (*e.g.*, PVP, PINC, etc.), peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (*see, e.g.*, U.S. Patent No. 5,922,687). Accordingly, peptides of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as Modified Vaccinia Ankara (MVA) (*e.g.*, Bavarian Noridic), vaccinia or fowlpox. For example, vaccinia virus is used as a vector to express nucleotide sequences that encode the peptides of the invention. Upon introduction into an acutely or chronically infected host or into a non-infected host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin).

BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, *e.g.* adeno and adeno-associated virus vectors, alpha virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, are apparent to those skilled in the art from the description herein.

[00248] In certain embodiments, components that induce T cell responses are combined with components that induce antibody responses to the target antigen of interest. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. Alternatively, a composition comprises a class I and/or class II epitope in accordance with the invention, along with a PADRE® molecule (Epimmune, San Diego, CA).

[00249] Compositions of the invention can comprise antigen presenting cells, such as dendritic cells. Antigen presenting cells, *e.g.*, dendritic cells, may be transfected, *e.g.*, with a polynucleotide such as a minigene construct in accordance with the invention, in order to elicit immune responses. The peptide can be bound to an HLA molecule on the antigen-resenting cell, whereby when an HLA-restricted cytotoxic T lymphocyte (CTL) is present, a receptor of the CTL binds to a complex of the HLA molecule and the peptide.

[00250] The compositions of the invention may also comprise antiviral drugs such as interferon-α, or immune adjuvants such as IL-12, GM-CSF, etc.

[00251] Compositions may comprise an HLA heavy chain, β_2-microglobulin, streptavidin, and/or biotin. The streptavidin may be fluorescently labeled. Compositions may comprise tetramers (*see e.g.*, U.S. Pat. No. 5,635,363; Science 274:94-96 (1996)). A tetramer composition comprising an HLA heavy chain, β_2-microglobulin, streptavidin, and biotin. The streptavidin may be fluorescently labeled. Compositions may also comprise dimers. A dimer composition comprises as MHC molecule and an Ig molecule (*see e.g.*, PNAS 95:7568-73 (1998)).

[00252] In some embodiments it may be desirable to include in the compositions of the invention at least one component which primes cytotoxic T lymphocytes. Lipids have been identified as agents capable of priming CTL *in vivo* against viral antigens. For example, palmitic acid residues can be attached to the ε-and α- amino groups of a lysine residue and then linked, *e.g.*, via one or more linking residues such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like, to an immunogenic peptide. The lipidated peptide can then be administered either directly in a micelle or particle, incorporated into a liposome, or

emulsified in an adjuvant, *e.g.*, incomplete Freund's adjuvant. A preferred composition comprises palmitic acid attached to ε- and α- amino groups of Lys, which is attached via linkage, *e.g.*, Ser-Ser, to the amino terminus of the peptide.

[00253] As another example of lipid priming of CTL responses, *E. coli* lipoproteins, such as tripalmitoyl-S-glycerylcysteinlyseryl-serine (P_3CSS) can be used to prime virus specific CTL when covalently attached to an appropriate peptide (*see, e.g.*, Deres, *et al., Nature* 342:561, 1989). Peptides of the invention can be coupled to P_3CSS, for example, and the lipopeptide administered to an individual to specifically prime a CTL response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be primed with P_3CSS-conjugated epitopes, two such compositions can be combined to more effectively elicit both humoral and cell-mediated responses.

[00254] Another preferred embodiment is a composition comprising one or more peptides of the invention emulsified in IFA.

[00255] Compositions of the invention may also comprise CTL and/or HTL peptides. Such CTL and HTL peptides can be modified by the addition of amino acids to the termini of a peptide to provide for ease of linking peptides one to another, for coupling to a carrier support or larger peptide, for modifying the physical or chemical properties of the peptide or oligopeptide, or the like. Amino acids such as tyrosine, cysteine, lysine, glutamic or aspartic acid, or naturally or unnaturally occuring amino acid residues, can be introduced at the carboxyl- or amino-terminus of the peptide or oligopeptide, particularly class I peptides. However, it is to be noted that modification at the carboxyl terminus of a CTL epitope may, in some cases, alter binding characteristics of the peptide. In addition, the peptide or oligopeptide sequences can differ from the natural sequence by being modified by terminal-NH_2 acylation, *e.g.*, by alkanoyl (C_1-C_{20}) or thioglycolyl acetylation, terminal-carboxyl amidation, *e.g.*, ammonia, methylamine, *etc.* In some instances these modifications may provide sites for linking to a support or other molecule. CTL and HTL epitopes may comprise additional amino acids, such as those described above including spacers.

[00256] A further embodiment of a composition in accordance with the invention is an antigen presenting cell that comprises one or more peptides in accordance with the invention. The antigen presenting cell can be a "professional" antigen presenting cell, such as a dendritic cell. The antigen presenting cell can comprise the peptide of the invention by any means known or to be determined in the art. Such means include pulsing

of dendritic cells with one or more individual peptides, by nucleic acid administration such as ballistic nucleic acid delivery or by other techniques in the art for administration of nucleic acids, including vector-based, *e.g.* viral vector, delivery of nucleic acids.

[00257] Compositions may comprise carriers. Carriers that can be used with compositions of the invention are well known in the art, and include, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza virus proteins, hepatitis B virus core protein, and the like.

[00258] The compositions (e.g. pharmaceutical compositions) can contain a physiologically tolerable diluent such as water, or a saline solution, preferably phosphate buffered saline. Additionally, as disclosed herein, CTL responses can be primed by conjugating peptides of the invention to lipids, such as tripalmitoyl-S-glyceryl-cysteinyl-seryl-serine (P_3CSS).

[00259] Compositions of the invention may be pharmaceutically acceptable compositions. Pharmaceutical compositions preferably contain an immunologically effective amount of one or more peptides and/or polynucleotides of the invention, and optionally one or more other components which are pharmaceutically acceptable. A preferred composition comprises one or more peptides of the invention and IFA. A more preferred composition of the invention comprises one or more peptides of the invention, one or more peptides, and IFA.

[00260] Upon immunization with a peptide and/or polynucleotide and/or composition in accordance with the invention, via injection (*e.g.*, SC, ID, IM), aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by an immune response comprising the production of antibodies, CTLs and/or HTLs specific for the desired antigen(s). Consequently, the host becomes at least partially immune to subsequent exposure to the infectious agent(s), or at least partially resistant to further development of infectious agent-bearing cells and thereby derives a prophylactic or therapeutic benefit.

[00261] Furthermore, the peptides, primers, and epitopes of the invention can be used in any desired immunization or administration regimen; e.g., as part of periodic vaccinations such as annual vaccinations as in the veterinary arts or as in periodic vaccinations as in the human medical arts, or as in a prime-boost regime wherein an inventive vector or recombinant is administered either before or after the administration of the same or of a different epitope of interest or recombinant or vector expressing such as a same or different epitope of interest (including an inventive recombinant or vector expressing such

as a same or different epitope of interest), see, e.g., U.S. Pat. Nos. 5,997,878; 6,130,066; 6,180,398; 6,267,965; and 6,348,450. An useful viral vector of the present invention is Modified Vaccinia Ankara (MVA) (*e.g.*, Bavarian Noridic (MVA-BN)).

[00262] Recent studies have indicated that a prime-boost protocol, whereby immunization with a poxvirus recombinant expressing a foreign gene product is followed by a boost using a purified subunit preparation form of that gene product, elicits an enhanced immune response relative to the response elicited with either product alone. Human volunteers immunized with a vaccinia recombinant expressing the HIV-1 envelope glycoprotein and boosted with purified HIV-1 envelope glycoprotein subunit preparation exhibit higher HIV-1 neutralizing antibody titers than individuals immunized with just the vaccinia recombinant or purified envelope glycoprotein alone (Graham *et al.*, *J. Infect. Dis.*, 167:533-537 (1993); Cooney *et al.*, *Proc. Natl. Acad. Sci. USA*, 90:1882-1886 (1993)). Humans immunized with two injections of an ALVAC-HIV-1 env recombinant (vCP125) failed to develop HIV specific antibodies. Boosting with purified rgp160 from a vaccinia virus recombinant resulted in detectable HIV-1 neutralizing antibodies. Furthermore, specific lymphocyte T cell proliferation to rgp160 was clearly increased by the boost with rgp160. Envelope specific cytotoxic lymphocyte activity was also detected with this vaccination regimen (Pialoux *et al.*, *AIDS Res. and Hum. Retroviruses*, 11:272-381 (1995)). Macaques immunized with a vaccinia recombinant expressing the simian immunodeficiency virus (SIV) envelope glycoprotein and boosted with SIV envelope glycoprotein from a baculovirus recombinant are protected against SIV challenge (Hu *et al.*, *AID Res. and Hum. Retroviruses*, 3:615-620 (1991); Hu *et al.*, *Science* 255:456-459 (1992)). In the same fashion, purified HCMVgB protein can be used in prime-boost protocols with NYVAC or ALVAC-gB recombinants.

[00263] In certain embodiments, the polynucleotides are complexed in a liposome preparation. Liposomal preparations for use in the instant invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. However, cationic liposomes are particularly preferred because a tight charge complex can be formed between the cationic liposome and the polyanionic nucleic acid. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Felgner *et al.*, *Proc. Natl. Acad. Sci. USA 84*:74137416 (1987), which is herein incorporated by reference); mRNA (Malone *et al.*, *Proc. Natl. Acad. Sci. USA 86*:60776081 (1989), which is herein incorporated by reference); and purified transcription factors (Debs *et al.*, *J. Biol.*

Chem. 265:1018910192 (1990), which is herein incorporated by reference), in functional form.

[00264] Cationic liposomes are readily available. For example, N-[12,3-dioleyloxy)-propyl]-N,N,N-triethylammonium (DOTMA) liposomes are particularly useful and are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, N.Y. (See, also, Felgner *et al.*, *Proc. Natl Acad. Sci. USA 84*:74137416 (1987)). Other commercially available liposomes include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boehringer).

[00265] Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g. PCT Publication No. WO 90/11092 for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes. Preparation of DOTMA liposomes is explained in the literature, see, e.g., P. Felgner *et al.*, *Proc. Natl. Acad. Sci. USA 84*:74137417. Similar methods can be used to prepare liposomes from other cationic lipid materials.

[00266] Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, Ala.), or can be easily prepared using readily available materials. Such materials include phosphatidyl, choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphoshatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

[00267] For example, commercially available dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), and dioleoylphosphatidyl ethanolamine (DOPE) can be used in various combinations to make conventional liposomes, with or without the addition of cholesterol. Thus, for example, DOPG/DOPC vesicles can be prepared by drying 50 mg each of DOPG and DOPC under a stream of nitrogen gas into a sonication vial. The sample is placed under a vacuum pump overnight and is hydrated the following day with deionized water. The sample is then sonicated for 2 hours in a capped vial, using a Heat Systems model 350 sonicator equipped with an inverted cup (bath type) probe at the maximum setting while the bath is circulated at 15EC. Alternatively, negatively charged vesicles can be prepared without sonication to produce multilamellar vesicles or by extrusion through nucleopore membranes to produce unilamellar vesicles of discrete size. Other methods are known and available to those of skill in the art.

[00268] The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs), with SUVs being preferred. The various liposome nucleic acid complexes are prepared using methods well known in the art. See, e.g., Straubinger *et al.*, *Methods of Immunology 101*:512527 (1983). For example, MLVs containing nucleic acid can be prepared by depositing a thin film of phospholipid on the walls of a glass tube and subsequently hydrating with a solution of the material to be encapsulated. SUVs are prepared by extended sonication of MLVs to produce a homogeneous population of unilamellar liposomes. The material to be entrapped is added to a suspension of preformed MLVs and then sonicated. When using liposomes containing cationic lipids, the dried lipid film is resuspended in an appropriate solution such as sterile water or an isotonic buffer solution such as 10 mM Tris/NaCl, sonicated, and then the preformed liposomes are mixed directly with the DNA. The liposome and DNA form a very stable complex due to binding of the positively charged liposomes to the cationic DNA. SUVs find use with small nucleic acid fragments. LUVs are prepared by a number of methods, well known in the art. Commonly used methods include Ca^{2+}-EDTA chelation (Papahadjopoulos *et al.*, *Biochim. Biophys. Acta 394*:483 (1975); Wilson *et al.*, *Cell 17*:77 (1979)); ether injection (Deamer, D. and Bangham, A.,. *Biochim. Biophys. Acta 443*:629 (1976); Ostro *et al.*, *Biochem. Biophys. Res. Commun. 76*:836 (1977); Fraley *et al.*, *Proc. Natl. Acad. Sci. USA 76*:3348 (1979)); detergent dialysis (Enoch, H. and Strittmatter, P., *Proc. Natl. Acad. Sci. USA 76*:145 (1979)); and reversephase evaporation (REV) (Fraley *et al.*, *J. Biol. Chem. 255*:10431 (1980); Szoka, F. and Papahadjopoulos, D., *Proc. Natl. Acad. Sci. USA 75*:145 (1978); SchaeferRidder *et al.*, *Science 215*:166 (1982)).

[00269] Generally, the ratio of DNA to liposomes will be from about 10:1 to about 1:10. Preferably, the ration will be from about 5:1 to about 1:5. More preferably, the ration will be about 3:1 to about 1:3. Still more preferably, the ratio will be about 1:1.

[00270] U.S. Patent No. 5,676,954 reports on the injection of genetic material, complexed with cationic liposome carriers, into mice. U.S. Patent Nos. 4,897,355, 4,946,787, 5,049,386, 5,459,127, 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication no. WO 94/9469 provide cationic lipids for use in transfecting DNA into cells and mammals. U.S. Patent Nos. 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication no. WO 94/9469 provide methods for delivering DNA-cationic lipid complexes to mammals.

Binding Affinity of Variants for HLA Molecules

[00271] As indicated herein, the large degree of HLA polymorphism is an important factor to be taken into account with the epitope-based approach to developing therapeutics and diagnostics. To address this factor, epitope selection encompassing identification of peptides capable of binding at high or intermediate affinity to multiple HLA molecules is preferably utilized, most preferably these epitopes bind at high or intermediate affinity to two or more allele-specific HLA molecules. However, in some embodiments, it is preferred that all epitopes in a given composition bind to the alleles of a single HLA supertype or a single HLA molecule.

[00272] Variants of the invention preferably include those that have an IC_{50} or binding affinity value for a class I HLA molecule(s) of 500 nM or better (*i.e.*, the value is ≤ 500 nM). In certain embodiments of the invention, peptides of interest have an IC_{50} or binding affinity value for a class I HLA molecule(s) of 200 nM or better. In certain embodiments of the invention, peptides of interest, such as A1 and A24 peptides, have an IC_{50} or binding affinity value for a class I HLA molecule(s) of 100 nM or better. If HTL epitopes are included, they preferably are HTL epitopes that have an IC_{50} or binding affinity value for class II HLA molecules of 1000 nM or better, (*i.e.*, the value is ≤ 1,000 nM). For example, peptide binding is assessed by testing the capacity of a candidate peptide to bind to a purified HLA molecule *in vitro*. Peptides exhibiting high or intermediate affinity are then considered for further analysis. Selected peptides are generally tested on other members of the supertype family. In preferred embodiments, peptides that exhibit cross-reactive binding are then used in cellular screening analyses or vaccines.

[00273] The relationship between binding affinity for HLA class I molecules and immunogenicity of discrete peptide epitopes on bound antigens was determined for the first time by inventors at Epimmune. As disclosed in greater detail herein, higher HLA binding affinity is correlated with greater immunogenicity.

[00274] Greater immunogenicity can be manifested in several different ways. Immunogenicity corresponds to whether an immune response is elicited at all, and to the vigor of any particular response, as well as to the extent of a population in which a response is elicited. For example, a peptide might elicit an immune response in a diverse array of the population, yet in no instance produce a vigorous response. In accordance

with these principles, close to 90% of high binding peptides have been found to elicit a response and thus be "immunogenic," as contrasted with about 50% of the peptides that bind with intermediate affinity. (See, e.g., Schaeffer *et al.* PNAS (1988)) High affinity-binding class I peptides generally have an affinity of less than or equal to 100 nM. Moreover, not only did peptides with higher binding affinity have an enhanced probability of generating an immune response, the generated response tended to be more vigorous than the response seen with weaker binding peptides. As a result, less peptide is required to elicit a similar biological effect if a high affinity binding peptide is used rather than a lower affinity one. Thus, in some preferred embodiments of the invention, high affinity binding epitopes are used.

[00275] The correlation between binding affinity and immunogenicity was analyzed by the present inventors by two different experimental approaches (*see, e.g.*, Sette, *et al., J. Immunol.* 153:5586-5592 (1994)). In the first approach, the immunogenicity of potential epitopes ranging in HLA binding affinity over a 10,000-fold range was analyzed in HLA-A*0201 transgenic mice. In the second approach, the antigenicity of approximately 100 different hepatitis B virus (HBV)-derived potential epitopes, all carrying A*0201 binding motifs, was assessed by using PBL from acute hepatitis patients. Pursuant to these approaches, it was determined that an affinity threshold value of approximately 500 nM (preferably 50 nM or less) determines the capacity of a peptide epitope to elicit a CTL response. These data are true for class I binding affinity measurements for naturally processed peptides and for synthesized T cell epitopes. These data also indicate the important role of determinant selection in the shaping of T cell responses (*see, e.g.*, Schaeffer *et al. Proc. Natl. Acad. Sci. USA* 86:4649-4653 (1989)).

[00276] An affinity threshold associated with immunogenicity in the context of HLA class II (*i.e.*, HLA DR) molecules has also been delineated (*see, e.g.*, Southwood *et al. J. Immunology* 160:3363-3373 (1998), and U.S. Patent No. 6,413,527, issued July 2, 2002). In order to define a biologically significant threshold of HLA class II binding affinity, a database of the binding affinities of 32 DR-restricted epitopes for their restricting element (*i.e.*, the HLA molecule that binds the epitope) was compiled. In approximately half of the cases (15 of 32 epitopes), DR restriction was associated with high binding affinities, *i.e.* binding affinity values of 100 nM or less. In the other half of the cases (16 of 32), DR restriction was associated with intermediate affinity (binding affinity values in the 100-1000 nM range). In only one of 32 cases was DR restriction associated with an IC_{50}

of 1000 nM or greater. Thus, 1000 nM is defined as an affinity threshold associated with immunogenicity in the context of DR molecules.

[00277] The binding affinity of peptides for HLA molecules can be determined as described in Example 1, below.

Enhancing Population Coverage of the Vaccine

[00278] The primary anchor residues of the HLA class I peptide epitope supermotifs and motifs are summarized in Tables 1-2. Allele-specific HLA molecules that are comprised by the various HLA class I supertypes are listed in Table 4. In some cases, patterns of amino acid residues are present in both a motif and a supermotif. The relationship of a particular motif and any related supermotif is indicated in the description of the individual motifs.

[00279] By inclusion of one or more epitopes from several motifs or supermotifs in a vaccine composition, enhanced population coverage for major global ethnicities can be obtained.

Assays to Detect T-Cell Responses

[00280] Once HLA binding peptides are identified, they can be tested for the ability to elicit a T-cell response. The preparation and evaluation of motif-bearing peptides are described, *e.g.*, in PCT publications WO 94/20127 and WO 94/03205. Briefly, peptides comprising epitopes from a particular antigen are synthesized and tested for their ability to bind to relevant HLA proteins. These assays may involve evaluation of peptide binding to purified HLA class I molecules in relation to the binding of a radioiodinated reference peptide. Alternatively, cells expressing empty class I molecules (*i.e.* cell surface HLA molecules that lack any bound peptide) may be evaluated for peptide binding by immunofluorescent staining and flow microfluorimetry. Other assays that may be used to evaluate peptide binding include peptide-dependent class I assembly assays and/or the inhibition of CTL recognition by peptide competition. Those peptides that bind to an HLA class I molecule, typically with an affinity of 500 nM or less, are further evaluated for their ability to serve as targets for CTLs derived from infected or immunized individuals, as well as for their capacity to induce primary *in vitro* or *in vivo* CTL

responses that can give rise to CTL populations capable of reacting with selected target cells associated with pathology.

[00281] Analogous assays are used for evaluation of HLA class II binding peptides. HLA class II motif-bearing peptides that are shown to bind, typically at an affinity of 1000 nM or less, are further evaluated for the ability to stimulate HTL responses.

[00282] Conventional assays utilized to detect T cell responses include proliferation assays, lymphokine secretion assays, direct cytotoxicity assays, and limiting dilution assays. For example, antigen-presenting cells that have been incubated with a peptide can be assayed for the ability to induce CTL responses in responder cell populations. Antigen-presenting cells can be normal cells such as peripheral blood mononuclear cells or dendritic cells. Alternatively, mutant, non-human mammalian cell lines that have been transfected with a human class I MHC gene, and that are deficient in their ability to load class I molecules with internally processed peptides, are used to evaluate the capacity of the peptide to induce *in vitro* primary CTL responses. Peripheral blood mononuclear cells (PBMCs) can be used as the source of CTL precursors. Antigen presenting cells are incubated with peptide, after which the peptide-loaded antigen-presenting cells are then incubated with the responder cell population under optimized culture conditions. Positive CTL activation can be determined by assaying the culture for the presence of CTLs that lyse radio-labeled target cells, either specific peptide-pulsed targets or target cells that express endogenously processed antigen from which the specific peptide was derived. Alternatively, the presence of epitope-specific CTLs can be determined by IFNγ *in situ* ELISA.

[00283] In an embodiment of the invention, directed to diagnostics, a method has been devised which allows direct quantification of antigen-specific T cells by staining with fluorescein-labelled HLA tetrameric complexes (Altman, J. D. *et al.*, *Proc. Natl. Acad. Sci. USA* 90:10330, 1993; Altman, J. D. *et al.*, *Science* 274:94, 1996). Other options include staining for intracellular lymphokines, and interferon release assays or ELISPOT assays. Tetramer staining, intracellular lymphokine staining and ELISPOT assays all appear to be at least 10-fold more sensitive than more conventional assays (Lalvani, A. *et al.*, *J. Exp. Med.* 186:859, 1997; Dunbar, P. R. *et al.*, *Curr. Biol.* 8:413, 1998; Murali-Krishna, K. *et al.*, *Immunity* 8:177, 1998). Additionally, DimerX technology can be used as a means of quantitation (*see, e.g.*, *Science* 274:94-99 (1996) and *Proc. Natl. Acad. Sci.* 95:7568-73 (1998)).

[00284] HTL activation may also be assessed using techniques known to those in the art, such as T cell proliferation or lymphokine secretion (*see, e.g.* Alexander *et al.*, *Immunity* 1:751-761, 1994).

[00285] Alternatively, immunization of HLA transgenic mice can be used to determine immunogenicity of peptide epitopes. Several transgenic mouse strains, *e.g.*, mice with human A2.1, A11 (which can additionally be used to analyze HLA-A3 epitopes), and B7 alleles have been characterized. Other transgenic mice strains (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. Moreover, HLA-DR1 and HLA-DR3 mouse models have been developed. In accordance with principles in the art, additional transgenic mouse models with other HLA alleles are generated as necessary.

[00286] Such mice can be immunized with peptides emulsified in Incomplete Freund's Adjuvant; thereafter any resulting T cells can be tested for their capacity to recognize target cells that have been peptide-pulsed or transfected with genes encoding the peptide of interest. CTL responses can be analyzed using cytotoxicity assays described above. Similarly, HTL responses can be analyzed using, *e.g.*, T cell proliferation or lymphokine secretion assays.

Minigenes

[00287] A number of different approaches are available which allow simultaneous delivery of multiple epitopes. Nucleic acids encoding multiple epitopes are a useful embodiment of the invention; discrete peptide epitopes or polyepitopic peptides can be encoded. The epitopes to be included in a minigene are preferably selected according to the guidelines set forth in the previous section. Examples of amino acid sequences that can be included in a minigene include: HLA class I epitopes, HLA class II epitopes, a ubiquitination signal sequence, and/or a targeting sequence such as an endoplasmic reticulum (ER) signal sequence to facilitate movement of the resulting peptide into the endoplasmic reticulum. Examples of minigene constructs are shown in Tables 23-28.

[00288] The use of multi-epitope minigenes is also described in, *e.g.*, co-pending applications U.S.S.N. 09/311,784, 09/894,018, 60/419,973, 60/415,463; Ishioka *et al.*, *J. Immunol.* 162:3915-3925, 1999; An, L. and Whitton, J. L., *J. Virol.* 71:2292, 1997; Thomson, S. A. *et al.*, *J. Immunol.* 157:822, 1996; Whitton, J. L. *et al.*, *J. Virol.* 67:348, 1993; Hanke, R. *et al.*, *Vaccine* 16:426, 1998. For example, a multi-epitope DNA plasmid

encoding nine dominant HLA-A*0201- and A11-restricted CTL epitopes derived from the polymerase, envelope, and core proteins of HBV and human immunodeficiency virus (HIV), a PADRE® universal helper T cell (HTL) epitope, and an endoplasmic reticulum-translocating signal sequence has been engineered. Immunization of HLA transgenic mice with this plasmid construct resulted in strong CTL induction responses against the nine CTL epitopes tested. This CTL response was similar to that observed with a lipopeptide of known immunogenicity in humans, and significantly greater than immunization using peptides in oil-based adjuvants. Moreover, the immunogenicity of DNA-encoded epitopes *in vitro* was also correlated with the *in vitro* responses of specific CTL lines against target cells transfected with the DNA plasmid. These data show that the minigene served: 1.) to generate a CTL response and 2.) to generate CTLs that recognized cells expressing the encoded epitopes. A similar approach can be used to develop minigenes encoding epitopes of an infectious agent.

[00289] For example, to create a DNA sequence encoding the selected epitopes (minigene) for expression in human cells, the amino acid sequences of the epitopes may be reverse translated. A human codon usage table can be used to guide the codon choice for each amino acid. These epitope-encoding DNA sequences may be directly adjoined, so that when translated, a continuous peptide sequence is created. However, to optimize expression and/or immunogenicity, additional elements can be incorporated into the minigene design such as spacer amino acid residues between epitopes. HLA presentation of CTL and HTL epitopes may be improved by including synthetic (*e.g.* poly-alanine) or naturally-occurring flanking sequences adjacent to the CTL or HTL epitopes; these larger peptides comprising the epitope(s) are within the scope of the invention. In one embodiment, spacer amino acid residues between one or more CTL and/or HTL epitopes are designed so as to minimize junctional epitopes that may result from the juxtaposition of 2 CTL and/or HTL epitopes.

[00290] The minigene sequence may be converted to DNA by assembling oligonucleotides that encode the plus and minus strands of the minigene. Overlapping oligonucleotides (30-100 bases long) may be synthesized, phosphorylated, purified and annealed under appropriate conditions using well known techniques. The ends of the oligonucleotides can be joined, for example, using T4 DNA ligase. This synthetic minigene, encoding the epitope peptide, can then be cloned into a desired expression vector.

[00291] Standard regulatory sequences well known to those of skill in the art are preferably included in the vector to ensure expression in the target cells. Several vector elements are desirable: a promoter with a downstream cloning site for minigene insertion; a polyadenylation signal for efficient transcription termination; an *E. coli* origin of replication; and an *E. coli* selectable marker (*e.g.* ampicillin or kanamycin resistance). Numerous promoters can be used for this purpose, *e.g.*, the human cytomegalovirus (hCMV) CMV-IE promoter. See, *e.g.*, U.S. Patent Nos. 5,580,859 and 5,589,466 for other suitable promoter sequences.

[00292] Optimized peptide expression and immunogenicity can be achieved by certain modifications to a minigene construct. For example, in some cases introns facilitate efficient gene expression, thus one or more synthetic or naturally-occurring introns can be incorporated into the transcribed region of the minigene. The inclusion of mRNA stabilization sequences and sequences for replication in mammalian cells may also be considered for increasing minigene expression.

[00293] Once an expression vector is selected, the minigene is cloned into the polylinker region downstream of the promoter. This plasmid is transformed into an appropriate bacterial strain, and DNA is prepared using standard techniques. The orientation and DNA sequence of the minigene, as well as all other elements included in the vector, are confirmed using restriction mapping, PCR and/or DNA sequence analysis. Bacterial cells harboring the correct plasmid can be stored as cell banks.

[00294] In addition, immunostimulatory sequences (ISSs or CpGs) appear to play a role in the immunogenicity of DNA vaccines. These sequences may be included in the vector, outside the minigene coding sequence to enhance immunogenicity.

[00295] In some embodiments, a bi-cistronic expression vector which allows production of both the minigene-encoded epitopes and a second protein (*e.g.*, one that modulates immunogenicity) can be used. Examples of proteins or polypeptides that, if co-expressed with epitopes, can enhance an immune response include cytokines (*e.g.*, IL-2, IL-12, GM-CSF), cytokine-inducing molecules (*e.g.*, LeIF), costimulatory molecules, or pan-DR binding proteins (PADRE®, Epimmune, San Diego, CA). Helper T cell (HTL) epitopes such as PADRE® molecules can be joined to intracellular targeting signals and expressed separately from expressed CTL epitopes. This can be done in order to direct HTL epitopes to a cell compartment different than that of the CTL epitopes, one that provides for more efficient entry of HTL epitopes into the HLA class II pathway, thereby

improving HTL induction. In contrast to HTL or CTL induction, specifically decreasing the immune response by co-expression of immunosuppressive molecules (*e.g.* TGF-β) may be beneficial in certain diseases.

[00296] Therapeutic quantities of plasmid DNA can be produced for example, by fermentation in *E. coli*, followed by purification. Aliquots from the working cell bank are used to inoculate growth medium, and are grown to saturation in shaker flasks or a bioreactor according to well known techniques. Plasmid DNA is purified using standard bioseparation technologies such as solid phase anion-exchange resins available, *e.g.*, from QIAGEN, Inc. (Valencia, California). If required, supercoiled DNA can be isolated from the open circular and linear forms using gel electrophoresis or other methods.

[00297] Purified plasmid DNA can be prepared for injection using a variety of formulations. The simplest of these is reconstitution of lyophilized DNA in sterile phosphate-buffer saline (PBS). This approach, known as "naked DNA," is currently being used for intramuscular (IM) administration in clinical trials. To maximize the immunotherapeutic effects of minigene vaccines, alternative methods of formulating purified plasmid DNA may be used. A variety of such methods have been described, and new techniques may become available. Cationic lipids, glycolipids, and fusogenic liposomes can also be used in the formulation (see, *e.g.*, WO 93/24640; Mannino & Gould-Fogerite, *BioTechniques* 6(7): 682 (1988); U.S. Patent No. 5,279,833; WO 91/06309; and Felgner, *et al., Proc. Nat'l Acad. Sci. USA* 84:7413 (1987). In addition, peptides and compounds referred to collectively as protective, interactive, non-condensing compounds (PINC) can also be complexed to purified plasmid DNA to influence variables such as stability, intramuscular dispersion, or trafficking to specific organs or cell types.

[00298] Known methods in the art can be used to enhance delivery and uptake of a polynucleotide *in vivo*. For example, the polynucleotide can be complexed to polyvinylpyrrolidone (PVP), to prolong the localized bioavailability of the polynucleotide, thereby enhancing uptake of the polynucleotide by the organisum (*see e.g.*, U.S. Patent No. 6,040,295; EP 0 465 529; WO 98/17814). PVP is a polyamide that is known to form complexes with a wide variety of substances, and is chemically and physiologically inert.

[00299] Target cell sensitization can be used as a functional assay of the expression and HLA class I presentation of minigene-encoded epitopes. For example, the plasmid DNA is introduced into a mammalian cell line that is a suitable target for standard CTL chromium release assays. The transfection method used will be dependent on the final

formulation, electroporation can be used for "naked" DNA, whereas cationic lipids or DNA:PVP compositions allow direct *in vitro* transfection. A plasmid expressing green fluorescent protein (GFP) can be co-transfected to allow enrichment of transfected cells using fluorescence activated cell sorting (FACS). The transfected cells are then chromium-51 (^{51}Cr) labeled and used as targets for epitope-specific CTLs. Cytolysis of the target cells, detected by ^{51}Cr release, indicates both the production and HLA presentation of, minigene-encoded CTL epitopes. Expression of HTL epitopes may be evaluated in an analogous manner using assays to assess HTL activity.

[00300] *In vivo* immunogenicity is a second approach for functional testing of minigene DNA formulations. Transgenic mice expressing appropriate human HLA proteins are immunized with the DNA product. The dose and route of administration are formulation dependent (*e.g.*, IM for DNA in PBS, intraperitoneal (IP) for lipid-complexed DNA). Eleven to twenty-one days after immunization, splenocytes are harvested and restimulated for one week in the presence of peptides encoding each epitope being tested. Thereafter, for CTLs, standard assays are conducted to determine if there is cytolysis of peptide-loaded, ^{51}Cr-labeled target cells. Once again, lysis of target cells that were exposed to epitopes corresponding to those in the minigene, demonstrates DNA vaccine function and induction of CTLs. Immunogenicity of HTL epitopes is evaluated in transgenic mice in an analogous manner.

[00301] Alternatively, the nucleic acids can be administered using ballistic delivery as described, for instance, in U.S. Patent No. 5,204,253. Using this technique, particles comprised solely of DNA are administered. In a further alternative embodiment for ballistic delivery, DNA can be adhered to particles, such as gold particles.

Vaccine Compositions

[00302] Vaccines that contain an immunologically effective amount of one or more peptides or polynucleotides of the invention are a further embodiment of the invention. The peptides can be delivered by various means or formulations, all collectively referred to as "vaccine" compositions. Such vaccine compositions, and/or modes of administration, can include, for example, naked DNA, DNA formulated with PVP, DNA in cationic lipid formulations; lipopeptides (*e.g.*,Vitiello, A. *et al.*, *J. Clin. Invest.* 95:341, 1995), DNA or peptides, encapsulated *e.g.*, in poly(DL-lactide-co-glycolide) ("PLG") microspheres (*see, e.g.*, Eldridge, *et al.*, *Molec. Immunol.* 28:287-294, 1991: Alonso *et al.*,

Vaccine 12:299-306, 1994; Jones *et al.*, *Vaccine* 13:675-681, 1995); peptide compositions contained in immune stimulating complexes (ISCOMS) (*see, e.g.*, Takahashi *et al.*, *Nature* 344:873-875, 1990; Hu *et al.*, *Clin Exp Immunol.* 113:235-243, 1998); multiple antigen peptide systems (MAPs) (*see e.g.*, Tam, J. P., *Proc. Natl. Acad. Sci. U.S.A.* 85:5409-5413, 1988; Tam, J.P., *J. Immunol. Methods* 196:17-32, 1996); viral, bacterial, or, fungal delivery vectors (Perkus, M. E. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 379, 1996; Chakrabarti, S. *et al.*, *Nature* 320:535, 1986; Hu, S. L. *et al.*, *Nature* 320:537, 1986; Kieny, M.-P. *et al.*, *AIDS Bio/Technology* 4:790, 1986; Top, F. H. *et al.*, *J. Infect. Dis.* 124:148, 1971; Chanda, P. K. *et al.*, *Virology* 175:535, 1990); particles of viral or synthetic origin (*e.g.*, Kofler, N. *et al.*, *J. Immunol. Methods.* 192:25, 1996; Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993; Falo, L. D., Jr. *et al.*, *Nature Med.* 7:649, 1995); adjuvants (*e.g.*, incomplete freund's advjuvant) (Warren, H. S., Vogel, F. R., and Chedid, L. A. *Annu. Rev. Immunol.* 4:369, 1986; Gupta, R. K. *et al.*, *Vaccine* 11:293, 1993); liposomes (Reddy, R. *et al.*, *J. Immunol.* 148:1585, 1992; Rock, K. L., *Immunol. Today* 17:131, 1996); or, particle-absorbed DNA (Ulmer, J. B. *et al.*, *Science* 259:1745, 1993; Robinson, H. L., Hunt, L. A., and Webster, R. G., *Vaccine* 11:957, 1993; Shiver, J. W. *et al.*, In: *Concepts in vaccine development*, Kaufmann, S. H. E., ed., p. 423, 1996; Cease, K. B., and Berzofsky, J. A., *Annu. Rev. Immunol.* 12:923, 1994 and Eldridge, J. H. *et al.*, *Sem. Hematol.* 30:16, 1993), etc. Toxin-targeted delivery technologies, also known as receptor mediated targeting, such as those of Avant Immunotherapeutics, Inc. (Needham, Massachusetts) or attached to a stress protein, e.g., HSP 96 (Stressgen Biotechnologies Corp., Victoria, BC, Canada) can also be used.

[00303] Vaccines of the invention comprise nucleic acid mediated modalities. DNA or RNA encoding one or more of the peptides of the invention can be administered to a patient. This approach is described, for instance, in Wolff *et. al.*, *Science* 247:1465 (1990) as well as U.S. Patent Nos. 5,580,859; 5,589,466; 5,804,566; 5,739,118; 5,736,524; 5,679,647; and, WO 98/04720. Examples of DNA-based delivery technologies include "naked DNA", facilitated (bupivicaine, polymers (*e.g.*, PVP), peptide-mediated) delivery, cationic lipid complexes, and particle-mediated ("gene gun") or pressure-mediated delivery (*see, e.g.*, U.S. Patent No. 5,922,687). Accordingly, peptide vaccines of the invention can be expressed by viral or bacterial vectors. Examples of expression vectors include attenuated viral hosts, such as vaccinia or fowlpox. For example, vaccinia virus is used as a vector to express nucleotide sequences that encode the peptides of the invention

(*e.g.*, MVA). Upon introduction into an acutely or chronically infected host or into a non-infected host, the recombinant vaccinia virus expresses the immunogenic peptide, and thereby elicits an immune response. Vaccinia vectors and methods useful in immunization protocols are described in, *e.g.*, U.S. Patent No. 4,722,848. Another vector is BCG (Bacille Calmette Guerin). BCG vectors are described in Stover *et al.*, *Nature* 351:456-460 (1991). A wide variety of other vectors useful for therapeutic administration or immunization of the peptides of the invention, *e.g.* adeno and adeno-associated virus vectors, alpha virus vectors, retroviral vectors, *Salmonella typhi* vectors, detoxified anthrax toxin vectors, and the like, are apparent to those skilled in the art from the description herein.

[00304] Furthermore, vaccines in accordance with the invention can comprise one or more peptides of the invention. Accordingly, a peptide can be present in a vaccine individually; alternatively, the peptide can exist as a homopolymer comprising multiple copies of the same peptide, or as a heteropolymer of various peptides. Polymers have the advantage of increased probability for immunological reaction and, where different peptide epitopes are used to make up the polymer, the ability to induce antibodies and/or T cells that react with different antigenic determinants of the antigen targeted for an immune response. The composition may be a naturally occurring region of an antigen or can be prepared, *e.g.*, recombinantly or by chemical synthesis.

[00305] Carriers that can be used with vaccines of the invention are well known in the art, and include, *e.g.*, thyroglobulin, albumins such as human serum albumin, tetanus toxoid, polyamino acids such as poly L-lysine, poly L-glutamic acid, influenza virus proteins, hepatitis B virus core protein, and the like. The vaccines can contain a physiologically tolerable diluent such as water, or a saline solution, preferably phosphate buffered saline. Generally, the vaccines also include an adjuvant. Adjuvants such as incomplete Freund's adjuvant, aluminum phosphate, aluminum hydroxide, or alum are examples of materials well known in the art. Additionally, as disclosed herein, CTL responses can be primed by conjugating peptides of the invention to lipids, such as tripalmitoyl-S-glyceryl-cysteinyl-seryl-serine (P_3CSS).

[00306] Upon immunization with a peptide composition in accordance with the invention, via injection (*e.g.*, SC, ID, IM), aerosol, oral, transdermal, transmucosal, intrapleural, intrathecal, or other suitable routes, the immune system of the host responds to the vaccine by producing antibodies, CTLs and/or HTLs specific for the desired antigen.

Consequently, the host becomes at least partially immune to subsequent exposure to the infectious agent, and thereby derives a prophylactic or therapeutic benefit.

[00307] In certain embodiments, components that induce T cell responses are combined with components that induce antibody responses to the target antigen of interest. A preferred embodiment of such a composition comprises class I and class II epitopes in accordance with the invention. Alternatively, a composition comprises a class I and/or class II epitope in accordance with the invention, along with a PADRE® molecule (Epimmune, San Diego, CA).

[00308] Vaccines of the invention can comprise antigen presenting cells, such as dendritic cells, as a vehicle to present peptides of the invention. For example, dendritic cells are transfected, *e.g.*, with a minigene construct in accordance with the invention, in order to elicit immune responses. Minigenes are discussed in greater detail in a following section. Vaccine compositions can be created *in vitro,* following dendritic cell mobilization and harvesting, whereby loading of dendritic cells occurs *in vitro*.

[00309] The vaccine compositions of the invention may also be used in combination with antiviral drugs such as interferon-α, or immune adjuvants such as IL-12, GM-CSF, etc.

[00310] Preferably, the following principles are utilized when selecting epitope(s) and/or analogs for inclusion in a vaccine, either peptide-based or nucleic acid-based formulations. Exemplary variants that may be utilized in a vaccine to treat or prevent infectious agent-mediated disease are set out in Tables 6-9 and Figures 1A-4. Each of the following principles can be balanced in order to make the selection. When multiple epitopes are to be used in a vaccine, the epitopes may be, but need not be, contiguous in sequence in the native antigen from which the epitopes are derived. Such multiple epitotes can refer to the order of epitopes within a peptide, or to the selection of epitopes that come from the same reagion, for use in either individual peptides or in a multi-epitopic peptide.

1.) Variants are selected which, upon administration, mimic immune responses that have been observed to be correlated with prevention or clearance of infectious disease. For HLA Class I, this generally includes 3-7 variants from at least one infectious agent or antigen thereof.

2.) Variants are selected that have the requisite binding affinity established to be correlated with immunogenicity: for HLA Class I an IC_{50} of 500 nM or less, or for Class II an IC_{50} of 1000 nM or less. For HLA Class I it is presently preferred to select a peptide having an IC_{50} of 200 nM or less, as this is believed to better correlate not only to induction of an immune response, but to *in vitro* tumor cell killing as well. For HLA A1 and A24, it is especially preferred to select a peptide having an IC_{50} of 100 nM or less.

3.) Supermotif bearing-variants, or a sufficient array of allele-specific motif-bearing variants, are selected to give broad population coverage. In general, it is preferable to have at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess the breadth of population coverage.

4.) Of particular relevance are "nested epitopes." Nested epitopes occur where at least two epitopes overlap in a given peptide sequence. For example, a nested epitope can be a fragment of an antigen from a region that contains multiple epitopes that are overleapping, or one epitope that is completely encompassed by another, *e.g.*, A2 peptides MAGE3.159 and MAGE3.160 are nested epitopes. A peptide comprising "transcendent nested epitopes" is a peptide that has both HLA class I and HLA class II epitopes in it. When providing nested epitopes, it is preferable to provide a sequence that has the greatest number of epitopes per provided sequence. Preferably, one avoids providing a peptide that is any longer than the amino terminus of the amino terminal epitope and the carboxyl terminus of the carboxyl terminal epitope in the peptide. When providing a sequence comprising nested epitopes, it is important to evaluate the sequence in order to insure that it does not have pathological or other deleterious biological properties; this is particularly relevant for vaccines directed to infectious organisms.

5.) If a protein with multiple epitopes or a polynucleotide (e.g., minigene) is created, an objective is to generate the smallest peptide that encompasses the epitopes of interest. This principle is similar, if not the same as that employed when selecting a peptide comprising nested epitopes. However, with an artificial peptide comprising multipe epitopes, the size minimization objective is balanced against the need to integrate any spacer sequences between epitopes in the polyepitopic protein. Spacer amino acid residues can be introduced to avoid junctional epitopes (an epitope recognized by the immune system, not present in the target antigen, and only created by the man-made juxtaposition of epitopes), or to facilitate cleavage between epitopes and thereby enhance epitope presentation. Junctional epitopes are generally to be avoided because the recipient may generate an immune response to that non-native epitope. Of particular concern is a junctional epitope that is a "dominant epitope." A dominant epitope may lead to such a zealous response that immune responses to other epitopes are diminished or suppressed.

[00311] The principles are the same, except junctional epitopes applies to the sequences surrounding the epitope. One must also take care with other sequences in construct to avoid immune response.

T CELL PRIMING MATERIALS

[00312] In some embodiments it may be desirable to include in the pharmaceutical compositions of the invention at least one component which primes cytotoxic T

lymphocytes. Lipids have been identified as agents capable of facilitating the priming *in vitro* CTL response against viral antigens. For example, palmitic acid residues can be attached to the ε-and α- amino groups of a lysine residue and then linked to an immunogenic peptide. One or more linking moieties can be used such as Gly, Gly-Gly-, Ser, Ser-Ser, or the like. The lipidated peptide can then be administered directly in a micelle or particle, incorporated into a liposome, or emulsified in an adjuvant, *e.g.*, incomplete Freund's adjuvant. A preferred immunogenic composition comprises palmitic acid attached to ε- and α- amino groups of Lys via a linking moiety, *e.g.*, Ser-Ser, added to the amino terminus of an immunogenic peptide.

[00313] In another embodiment of lipid-facilitated priming of CTL responses, *E. coli* lipoproteins, such as tripalmitoyl-S-glyceryl-cysteinyl-seryl-serine (P_3CSS) can be used to prime CTL when covalently attached to an appropriate peptide. (*See, e.g.*, Deres, *et al.*, *Nature* 342:561, 1989). Thus, peptides of the invention can be coupled to P_3CSS, and the lipopeptide administered to an individual to specifically prime a CTL response to the target antigen. Moreover, because the induction of neutralizing antibodies can also be primed with P_3CSS-conjugated epitopes, two such compositions can be combined to elicit both humoral and cell-mediated responses.

DENDRITIC CELLS PULSED WITH CTL AND/OR HTL PEPTIDES

[00314] An embodiment of a vaccine composition in accordance with the invention comprises *ex vivo* administration of a cocktail of epitope-bearing peptides to PBMC, or isolated DC therefrom, from the patient's blood. A pharmaceutical to facilitate harvesting of DC can be used, such as Progenipoietin™ (Monsanto, St. Louis, MO) or GM-CSF/IL-4. After pulsing the DC with peptides and prior to reinfusion into patients, the DC are washed to remove unbound peptides. In this embodiment, a vaccine comprises peptide-pulsed DCs which present the pulsed peptide epitopes in HLA molecules on their surfaces.

[00315] The DC can be pulsed *ex vivo* with a cocktail of peptides, some of which stimulate CTL responses to one or more antigens of interest, *e.g.*, antigens from infectious agents such as HIV env, HIV pol, HIV gag, HIV vpu, HBV and/or the antigens in Tables 11-22, or otherwise described herein or know in the art. Optionally, a helper T cell (HTL) peptide such as PADRE®, can be included to facilitate the CTL response. Thus, a vaccine in accordance with the invention comprising epitopes from an infectious agent is used to

treat or prevent disease mediated by these agents in patients. A vaccine can be used prior to, during, or following other therapies including, for example, antibiotic therepy, anti-viral therapy (e.g., highly active antiretroviral therapy (HAART) in the case of HIV-AIDS), antibody therapy, cancer therapy, and adjunct thereapy, whereupon the vaccine provides descreased morbidity, increased disease free survival and overall survival in recipients.

DIAGNOSTIC AND PROGNOSTIC USES

[00316] In one embodiment of the invention, HLA class I and class II binding peptides can be used as reagents to evaluate an immune response. Preferably, the following principles are utilized when selecting a variant(s) for diagnostic, prognostic and similar uses. Potential principles include having the binding affinities described earlier, and/or matching the HLA-motif/supermotif of a peptide with the HLA-type of a patient.

[00317] The evaluated immune response can be induced by any immunogen. For example, the immunogen may result in the production of antigen-specific CTLs or HTLs that recognize the peptide epitope(s) employed as the reagent. Thus, a peptide of the invention may or may not be used as the immunogen. Assay systems that can be used for such analyses include tetramer-based protocols (*e.g.*, DimerX technology (*see, e.g., Science* 274:94-99 (1996) and *Proc. Natl. Acad. Sci.* 95:7568-73 (1998)), staining for intracellular lymphokines, interferon release assays, or ELISPOT assays.

[00318] For example, following exposure to a putative immunogen, a peptide of the invention can be used in a tetramer staining assay to assess peripheral blood mononuclear cells for the presence of any antigen-specific CTLs. The HLA-tetrameric complex is used to directly visualize antigen-specific CTLs and thereby determine the frequency of such antigen-specific CTLs in a sample of peripheral blood mononuclear cells (*see, e.g.*, Ogg *et al.*, *Science* 279:2103-2106, 1998; and Altman *et al.*, *Science* 174:94-96, 1996).

[00319] A tetramer reagent comprising a peptide of the invention is generated as follows: A peptide that binds to an HLA molecule is refolded in the presence of the corresponding HLA heavy chain and β_2-microglobulin to generate a trimolecular complex. The complex is biotinylated at the carboxyl terminal end of the HLA heavy chain, at a site that was previously engineered into the protein. Tetramer formation is then induced by adding streptavidin. When fluorescently labeled streptavidin is used, the tetrameric complex is

used to stain antigen-specific cells. The labeled cells are then readily identified, *e.g.*, by flow cytometry. Such procedures are used for diagnostic or prognostic purposes; the cells identified by the procedure can be used for therapeutic purposes.

[00320] Peptides of the invention are also used as reagents to evaluate immune recall responses. (*see, e.g.*, Bertoni *et al.*, *J. Clin. Invest.* 100:503-513, 1997 and Penna *et al.*, *J. Exp. Med.* 174:1565-1570, 1991.) For example, a PBMC sample from an individual expressing a disease-associated antigen (e.g. an antigen from an infectious agent) can be analyzed for the presence of antigen-specific CTLs or HTLs using specific peptides. A blood sample containing mononuclear cells may be evaluated by cultivating the PBMCs and stimulating the cells with a peptide of the invention. After an appropriate cultivation period, the expanded cell population may be analyzed, for example, for CTL or for HTL activity.

[00321] Thus, the peptides can be used to evaluate the efficacy of a vaccine. PBMCs obtained from a patient vaccinated with an immunogen may be analyzed by methods such as those described herein. The patient is HLA typed, and peptide epitopes that are bound by the HLA molecule(s) present in that patient are selected for analysis. The immunogenicity of the vaccine is indicated by the presence of CTLs and/or HTLs directed to epitopes present in the vaccine.

[00322] The peptides of the invention may also be used to make antibodies, using techniques well known in the art (see, *e.g.* *CURRENT PROTOCOLS IN IMMUNOLOGY*, Wiley/Greene, NY; and *Antibodies A Laboratory Manual Harlow*, Harlow and Lane, Cold Spring Harbor Laboratory Press, 1989). Such antibodies are useful as reagents to determine the presence of disease-associated antigens. Antibodies in this category include those that recognize a peptide when bound by an HLA molecule, *i.e.*, antibodies that bind to a peptide-MHC complex.

ADMINISTRATION FOR THERAPEUTIC OR PROPHYLACTIC PURPOSES

[00323] The peptides and polynucleotides of the present invention, including cells and compositions comprising them, are useful for administration to mammals, particularly humans, to treat and/or prevent infection by an infectious agent such as HIV, HBV, HCV, HPV, *Plasmodium falciparum* and other agents described herein or known in the art. Vaccine compositions containing the peptides of the invention are administered to a

patient infected with a particular infectious agent or to an individual susceptible to, or otherwise at risk for, infection with such an agent to elicit an immune response against antigens of that agent and thus enhance the patient's own immune response capabilities. Where susceptible individuals are identified prior to infection, the composition can be targeted to them, thus minimizing the need for administration to a larger population.

[00324] In therapeutic applications, peptide and/or nucleic acid compositions are administered to a patient in an amount sufficient to elicit an effective immune response to the infectious agent antigen and to thereby cure, arrest or slow symptoms and/or complications. An amount adequate to accomplish this is defined as "therapeutically effective dose." Amounts effective for this use will depend on, *e.g.*, the particular composition administered, the manner of administration, the stage and severity of the disease being treated, the weight and general state of health of the patient, and the judgment of the prescribing physician.

[00325] The vaccine compositions of the invention can be used purely as prophylactic agents. Generally the dosage for an initial prophylactic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1000 μg of peptide and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg of peptide. Dosage values for a human typically range from about 500 μg to about 50,000 μg of peptide per 70 kilogram patient. This is followed by boosting dosages of between about 1.0 μg to about 50,000 μg of peptide, administered at defined intervals from about four weeks to six months after the initial administration of vaccine. The immunogenicity of the vaccine may be assessed by measuring the specific activity of CTL and HTL obtained from a sample of the patient's blood.

[00326] As noted above, peptides comprising CTL and/or HTL epitopes of the invention induce immune responses when presented by HLA molecules and contacted with a CTL or HTL specific for an epitope comprised by the peptide. The manner in which the peptide is contacted with the CTL or HTL is not critical to the invention. For instance, the peptide can be contacted with the CTL or HTL either *in vitro* or *in vivo*. If the contacting occurs *in vivo*, peptide can be administered directly, or in other forms/vehicles, *e.g.*, DNA vectors encoding one or more peptides, viral vectors encoding the peptide(s), liposomes, antigen presenting cells such as dendritic cells, and the like.

[00327] Accordingly, for pharmaceutical compositions of the invention in the form of peptides or polypeptides, the peptides or polypeptides can be administered directly. Alternatively, the peptide/polypeptides can be administered indirectly presented on APCs, or as DNA encoding them. Furthermore, the peptides or DNA encoding them can be administered individually or as fusions of one or more peptide sequences.

[00328] For therapeutic use, administration should generally begin at the first diagnosis of infectious agent-related disease. This is followed by boosting doses at least until symptoms are

substantially abated and for a period thereafter. In chronic disease states, loading doses followed by boosting doses may be required.

[00329] The dosage for an initial therapeutic immunization generally occurs in a unit dosage range where the lower value is about 1, 5, 50, 500, or 1,000 μg of peptide and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg of peptide. Dosage values for a human typically range from about 500 μg to about 50,000 μg of peptide per 70 kilogram patient. Boosting dosages of between about 1.0 μg to about 50,000 μg of peptide, administered pursuant to a boosting regimen over weeks to months, can be administered depending upon the patient's response and condition. Patient response can be determined by measuring the specific activity of CTL and HTL obtained from the patient's blood.

[00330] In certain embodiments, peptides and compositions of the present invention are used in serious disease states. In such cases, as a result of the minimal amounts of extraneous substances and the relative nontoxic nature of the peptides, it is possible and may be desirable to administer substantial excesses of these peptide compositions relative to these stated dosage amounts.

[00331] For treatment of chronic disease, a representative dose is in the range disclosed above, namely where the lower value is about 1, 5, 50, 500, or 1,000 μg of peptide and the higher value is about 10,000; 20,000; 30,000; or 50,000 μg of peptide, preferably from about 500 μg to about 50,000 μg of peptide per 70 kilogram patient. Initial doses followed by boosting doses at established intervals, *e.g.*, from four weeks to six months, may be required, possibly for a prolonged period of time to effectively immunize an individual. In the case of chronic disease, administration should continue until at least clinical symptoms or laboratory tests indicate that the disease has been eliminated or substantially abated, and for a follow-up period thereafter. The dosages, routes of administration, and dose schedules are adjusted in accordance with methodologies known in the art.

[00332] The pharmaceutical compositions for therapeutic treatment are intended for parenteral, topical, oral, intrathecal, or local administration. Preferably, the pharmaceutical compositions are administered parentally, *e.g.*, intravenously, subcutaneously, intradermally, or intramuscularly.

[00333] Thus, in a preferred embodiment the invention provides compositions for parenteral administration which comprise a solution of the immunogenic peptides dissolved or suspended in an acceptable carrier, preferably an aqueous carrier. A variety of aqueous carriers may be used, *e.g.*, water, buffered water, 0.8% saline, 0.3% glycine, hyaluronic acid and the like. These compositions may be sterilized by conventional, well known sterilization techniques, or may be sterile filtered. The resulting aqueous solutions may be packaged for use as is, or lyophilized, the lyophilized preparation being combined with a sterile solution prior to administration. The compositions may contain pharmaceutically acceptable auxiliary substances or pharmaceutical excipients as may be required to approximate physiological conditions, such as pH-adjusting and

buffering agents, tonicity adjusting agents, wetting agents, preservatives, and the like, for example, sodium acetate, sodium lactate, sodium chloride, potassium chloride, calcium chloride, sorbitan monolaurate, triethanolamine oleate, *etc.*

[00334] The concentration of peptides of the invention in the pharmaceutical formulations can vary widely, *i.e.*, from less than about 0.1%, usually at or at least about 2% to as much as 20% to 50% or more by weight, and will be selected primarily by fluid volumes, viscosities, *etc.*, in accordance with the particular mode of administration selected.

[00335] A human unit dose form of the peptide composition is typically included in a pharmaceutical composition that also comprises a human unit dose of an acceptable carrier, preferably an aqueous carrier, and is administered in a volume of fluid that is known by those of skill in the art to be used for administration of such compositions to humans (*see*, *e.g.*, *Remington's Pharmaceutical Sciences*, 17th Edition, A. Gennaro, Editor, Mack Publishing Co., Easton, Pennsylvania, 1985).

[00336] The peptides of the invention can also be administered via liposomes, which serve to target the peptides to a particular tissue, such as lymphoid tissue, or to target selectively to infected cells, as well as to increase the half-life of the peptide composition. Liposomes include emulsions, foams, micelles, insoluble monolayers, liquid crystals, phospholipid dispersions, lamellar layers and the like. In these preparations, the peptide to be delivered is incorporated as part of a liposome, alone or in conjunction with a molecule which binds to a receptor prevalent among lymphoid cells (such as monoclonal antibodies which bind to the CD45 antigen) or with other therapeutic or immunogenic compositions. Thus, liposomes either filled or decorated with a desired peptide of the invention can be directed to the site of lymphoid cells, where the liposomes then deliver the peptide compositions. Liposomes for use in accordance with the invention are formed from standard vesicle-forming lipids, which generally include neutral and negatively charged phospholipids and a sterol, such as cholesterol. The selection of lipids is generally guided by consideration of, *e.g.*, liposome size, acid lability and stability of the liposomes in the blood stream. A variety of methods are available for preparing liposomes, as described in, *e.g.*, Szoka, *et al.*, *Ann. Rev. Biophys. Bioeng.* 9:467 (1980), and U.S. Patent Nos. 4,235,871, 4,501,728, 4,837,028, and 5,019,369.

[00337] For targeting compositions of the invention to cells of the immune system, a ligand can be incorporated into the liposome, *e.g.*, antibodies or fragments thereof specific for cell surface determinants of the desired immune system cells. A liposome suspension containing a peptide may be administered intravenously, locally, topically, *etc.* in a dose which varies according to, *inter alia*, the manner of administration, the peptide being delivered, and the stage of the disease being treated.

[00338] For solid compositions, conventional nontoxic solid carriers may be used which include, for example, pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharin, talcum, cellulose, glucose, sucrose, magnesium carbonate, and the like. For oral administration, a pharmaceutically acceptable nontoxic composition is formed by incorporating any of the normally employed excipients, such as those carriers previously listed, and generally 10-95% of active ingredient, that is, one or more peptides of the invention, often at a concentration of 25%-75%.

[00339] For aerosol administration, the immunogenic peptides are preferably supplied in finely divided form, along with a surfactant and propellant. Typical percentages of peptides are 0.01%-20% by weight, often 1%-10%. The surfactant must, of course, be pharmaceutically acceptable, and preferably soluble in the propellant. Representative of such agents are the esters or partial esters of fatty acids containing from 6 to 22 carbon atoms, such as caproic, octanoic, lauric, palmitic, stearic, linoleic, linolenic, olesteric and oleic acids with an aliphatic polyhydric alcohol or its cyclic anhydride. Mixed esters, such as mixed or natural glycerides may be employed. The surfactant may constitute 0.1%-20% by weight of the composition, preferably 0.25-5%. The balance of the composition is ordinarily propellant, although an atomizer may be used in which no propellant is necessary and other percentages are adjusted accordingly. A carrier can also be included, *e.g.*, lecithin for intranasal delivery.

[00340] Antigenic peptides of the invention have been used to elicit a CTL and/or HTL response *ex vivo*, as well. The resulting CTLs or HTLs can be used to treat chronic infections, or tumors in patients that do not respond to other conventional forms of therapy, or who do not respond to a therapeutic peptide or nucleic acid vaccine in accordance with the invention. *Ex vivo* CTL or HTL responses to a particular antigen (infectious or tumor-associated) are induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of antigen-presenting cells (APC), such as dendritic cells, and the appropriate immunogenic peptide. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy (CTL) or facilitate destruction (HTL) of their specific target cell (an infected cell or a tumor cell).

KITS

[00341] The peptide and nucleic acid compositions of this invention can be provided in kit form together with instructions for vaccine administration. Typically the kit would include desired composition(s) of the invention in a container, preferably in unit dosage form and instructions for administration. For example , a kit would include an APC, such

as a dendritic cell, previously exposed to and now presenting peptides of the invention in a container, preferably in unit dosage form together with instructions for administration. An alternative kit would include a minigene construct with desired nucleic acids of the invention in a container, preferably in unit dosage form together with instructions for administration. Lymphokines such as IL-2 or IL-12 may also be included in the kit. Other kit components that may also be desirable include, for example, a sterile syringe, booster dosages, and other desired excipients.

[00342] The invention will be described in greater detail by way of specific examples. The following examples are offered for illustrative purposes, and are not intended to limit the invention in any manner. Those of skill in the art will readily recognize a variety of non-critical parameters that can be changed or modified to yield alternative embodiments in accordance with the invention.

EXAMPLES

EXAMPLE 1. HLA CLASS I AND CLASS II BINDING ASSAYS

[00343] The following example of peptide binding to HLA molecules demonstrates quantification of binding affinities of HLA class I and class II peptides. Binding assays can be performed with peptides that are either motif-bearing or not motif-bearing.

[00344] Cell lysates were prepared and HLA molecules purified in accordance with disclosed protocols (Sidney *et al.*, *Current Protocols in Immunology* 18.3.1 (1998); Sidney, *et al.*, *J. Immunol.* 154:247 (1995); Sette, *et al.*, *Mol. Immunol.* 31:813 (1994)). The cell lines used as sources of HLA molecules and the antibodies used for the extraction of the HLA molecules from the cell lysates are also described in these publications and are well known in the art.

[00345] Epstein-Barr virus (EBV)-transformed homozygous cell lines, fibroblasts, CIR, or 721.221-transfectants were used as sources of HLA class I molecules. These cells were cultured in RPMI 1640 medium supplemented with 2mM L-glutamine (GIBCO, Grand Island, NY), 50μM 2-ME, 100μg/ml of streptomycin, 100U/ml of penicillin (Irvine Scientific) and 10% heat-inactivated FCS (Irvine Scientific, Santa Ana, CA).

[00346] Cell lysates were prepared as follows. Briefly, cells were lysed at a concentration of 10^8 cells/ml in 50 mM Tris-HCl, pH 8.5, containing 1% Nonidet P-40 (Fluka

Biochemika, Buchs, Switzerland), 150 mM NaCl, 5 mM EDTA, and 2 mM PMSF. Lysates were cleared of debris and nuclei by centrifugation at 15,000 x g for 30min.

[00347] HLA molecules were purified from lysates by affinity chromatography. Lysates were passed twice through two pre-columns of inactivated Sepharose CL4-B and protein A-Sepharose. Next, the lysate was passed over a column of Sepharose CL-4B beads coupled to an appropriate antibody. The anti-HLA column was then washed with 10-column volumes of 10mM Tris-HCL, pH 8.0, in 1% NP-40, PBS, 2-column volumes of PBS, and 2-column volumes of PBS containing 0.4% n-octylglucoside. Finally, MHC molecules were eluted with 50mM diethylamine in 0.15M NaCl containing 0.4% n-octylglucoside, pH 11.5. A 1/25 volume of 2.0M Tris, pH 6.8, was added to the eluate to reduce the pH to ~8.0. Eluates were then concentrated by centrifugation in Centriprep 30 concentrators at 2000 rpm (Amicon, Beverly, MA). Protein content was evaluated by a BCA protein assay (Pierce Chemical Co., Rockford, IL) and confirmed by SDS-PAGE.

[00348] A detailed description of the protocol utilized to measure the binding of peptides to Class I and Class II MHC has been published (Sette *et al.*, *Mol. Immunol.* 31:813, 1994; Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998). Briefly, purified MHC molecules (5 to 500nM) were incubated with various unlabeled peptide inhibitors and 1-10nM ^{125}I-radiolabeled probe peptides for 48h in PBS containing 0.05% Nonidet P-40 (NP40) (or 20% w/v digitonin for H-2 IA assays) in the presence of a protease inhibitor cocktail. The final concentrations of protease inhibitors (each from CalBioChem, La Jolla, CA) were 1 mM PMSF, 1.3 nM 1.10 phenanthroline, 73 μM pepstatin A, 8mM EDTA, 6mM N-ethylmaleimide (for Class II assays), and 200 μM N alpha-p-tosyl-L-lysine chloromethyl ketone (TLCK). All assays were performed at pH 7.0 with the exception of DRB1*0301, which was performed at pH 4.5, and DRB1*1601 (DR2w21β_1) and DRB4*0101 (DRw53), which were performed at pH 5.0. pH was adjusted as described elsewhere (*see* Sidney *et al.*, in *Current Protocols in Immunology*, Margulies, Ed., John Wiley & Sons, New York, Section 18.3, 1998).

[00349] Following incubation, MHC-peptide complexes were separated from free peptide by gel filtration on 7.8 mm x 15 cm TSK200 columns (TosoHaas 16215, Montgomeryville, PA), eluted at 1.2 mls/min with PBS pH 6.5 containing 0.5% NP40 and 0.1% NaN_3. Because the large size of the radiolabeled peptide used for the DRB1*1501 (DR2w2β_1) assay makes separation of bound from unbound peaks more difficult under these conditions, all DRB1*1501 (DR2w2β_1) assays were performed using a 7.8mm x

30cm TSK2000 column eluted at 0.6 mls/min. The eluate from the TSK columns was passed through a Beckman 170 radioisotope detector, and radioactivity was plotted and integrated using a Hewlett-Packard 3396A integrator, and the fraction of peptide bound was determined.

[00350] Radiolabeled peptides were iodinated using the chloramine-T method. Representative radiolabeled probe peptides utilized in each assay, and its assay specific IC_{50} nM, are known in the art. Typically, in preliminary experiments, each MHC preparation was titered in the presence of fixed amounts of radiolabeled peptides to determine the concentration of HLA molecules necessary to bind 10-20% of the total radioactivity. All subsequent inhibition and direct binding assays were performed using these HLA concentrations.

[00351] Since under these conditions [label]<[HLA] and $IC_{50} \geq$[HLA], the measured IC_{50} values are reasonable approximations of the true K_D values. Peptide inhibitors are typically tested at concentrations ranging from 120 μg/ml to 1.2 ng/ml, and are tested in two to four completely independent experiments. To allow comparison of the data obtained in different experiments, a relative binding figure is calculated for each peptide by dividing the IC_{50} of a positive control for inhibition by the IC_{50} for each tested peptide (typically unlabeled versions of the radiolabeled probe peptide). For inter-experiment comparisons, relative binding values are compiled. These values can subsequently be converted back into IC_{50} nM values by dividing the IC_{50} nM of the positive controls for inhibition by the relative binding of the peptide of interest. This method of data compilation has proven to be the most accurate and consistent for comparing peptides that have been tested on different days, or with different lots of purified MHC.

[00352] Because the antibody used for HLA-DR purification (LB3.1) is α-chain specific, β_1 molecules are not separated from β_3 (and/or β_4 and β_5) molecules. The β_1 specificity of the binding assay is obvious in the cases of DRB1*0101 (DR1), DRB1*0802 (DR8w2), and DRB1*0803 (DR8w3), where no β_3 is expressed. It has also been demonstrated for DRB1*0301 (DR3) and DRB3*0101 (DR52a), DRB1*0401 (DR4w4), DRB1*0404 (DR4w14), DRB1*0405 (DR4w15), DRB1*1101 (DR5), DRB1*1201 (DR5w12), DRB1*1302 (DR6w19) and DRB1*0701 (DR7). The problem of β chain specificity for DRB1*1501 (DR2w2β_1), DRB5*0101 (DR2w2β_2), DRB1*1601 (DR2w21β_1), DRB5*0201 (DR51Dw21), and DRB4*0101 (DRw53) assays is circumvented by the use

of fibroblasts. Development and validation of assays with regard to DRβ molecule specificity have been described previously (*see, e.g.*, Southwood *et al.*, *J. Immunol.* 160:3363-3373, 1998).

[00353] Binding assays as outlined above may be used to analyze supermotif and/or motif-bearing epitopes.

EXAMPLE 2. RECOGNITION OF VARIANT PEPTIDES BY CTL DERIVED FROM DNA IMMUNIZATION

[00354] Variants corresponding to five HLA-A2 and –A3 restricted epitopes from 167 HIV varianst were identified and synthesized. These represented all the complete sequences in the Los Alamos database at the time (116 strains), as well as 51 complete clade C sequences from Botswana, and included 22 subtype B and 62 subtype C sequences. These peptides were then characterized with regard to MHC binding, variant distribution, and immunogenicity. To measure immunogenicity, HLA-A2/K^b or HLA-A11/K^b transgenic mice were immunized with the epitopes encoded in a DNA based format (). Eleven days after immunization, splenocytes were restimulated with either the epitope corresponding to the epitope encoded by the DNA (parent) or each of the variant peptides. After 6 days in culture, IFN-γ secretion was measured in response to the peptide used to stimulate each culture.

[00355] The data for these epitopes are shown in Figure 1. The HLA-A2-restricted epitope corresponding to the Env 134 epitope (KLTPLCVTL; Figure 1A) used as the immunogen was the form observed most often (134/167). All single anchor variants were recognized to approximately the same extent as the parent peptide. Many of the single non-anchor variants (9/13) were also recognized within 10-fold of the parent peptide. Conservative substitutions (R and Q for K; see Table 4) at position 1 (P1) were tolerated, while the non-conservative substitution (E for K; see Table 4) lowered binding and eliminated recognition. Three P4 variants were observed. Two of these (F or S for P) were recognized within 10-fold of the recognition of the parent peptide, while one substitution (Q for P) completely eliminated recognition. The binding for these peptides was not significantly different from the parent peptide, indicating that this residue may be involved in TCR recognition. Both the conservative (F for L) and non-conservative (R for L) substitutions

seen at P5 completely abrogated recognition, indicating that this residue is important in TCR recognition. Finally, one substitution at P8 (I for V), and four substitutions at P9 show little effect on recognition. None of the variants with multiple substitutions were recognized, although this may be due to the poor binding of these peptides.

[00356] The Gag 386 sequence utilized as the immunogen was the second most common form (VLAEAMSQV), present in 54 strains (Figure 1B). The most prevalent variant, differing by a single tolerated C terminal anchor residue (V to A; 67 strains), was recognized equally to the parent epitope by CTL raised against the parent, as were the remaining single-anchor variants. Single substitutions were also tolerated at the non-anchor positions, P1 (I for V) and P8 (R, K, or H for Q). Only the P7 variant (G for S), probably a TCR contact residue, was not recognized.

[00357] Many of the multiple variants for Gag 386 were also recognized by CTL raised against the parent peptide. All the variants with multiple changes combined a change of V to A or T at the C terminus with 1-3 additional substitutions. Two variants with N terminal changes (V to A or I) were observed. The non-conservative A substitution was not recognized, while the conservative I substitution was. A double variant with a conservative substitution at P3 (A to G) was not recognized, implicating P3 in TCR recognition. Double variants with conservative changes at position 8 (Q to R, K, or H) were not well recognized, although the variants with single changes at the same positions were recognized. The variant combining a non-conservative A residue at position 8 with A at the C terminus was recognized as well as the parent. Equally surprising was the observation that all the variants with 3 or 4 substitutions were recognized within 10-fold of the parent peptide.

[00358] The parent form of the HLA-A2-restricted epitope, Vpr 62 (RILQQLLFI; Figure 1C) was the most common form observed (86/167). Seven well-tolerated single anchor substitutions, 4 P2 and 3 C terminal, were also observed, accounting for most of the remaining variants (47/167). Single substitutions were, in general, also well tolerated. The single exception was the non-conservative substitution (P for L) at P6, while an M for L substitution at the same site was well tolerated. Binding was not affected for either variant, indicating that the reduction in activity is due to a change in a contact residue. Most variants with multiple changes also showed recognition to approximately the same extent as the parent. Several variants however did show reduced recognition. The variant with changes at both anchors (I to T at P2 and I to T at P9) had reduced binding (IC_{50} of 9700),

and recognition of the peptide was reduced, although not lost completely. Two variants with Q to H changes at P5, in combination with anchor residue changes (I to M at P2 and I to A at P9), exhibited greatly reduced recognition although binding was not affected. Other changes at P5 (Q to R or L at P5) reduced recognition only slightly.

[00359] The HLA-A3/11-restricted epitope, Pol 98 (Figure 1D), represented the most diverse epitope in terms of the number of variant epitopes identified. The peptide encoded in the DNA was represented in only 18 out of 167 strains. Approximately a third of the peptides identified at that position (49 out of 167) did not have recognizable A3/A11 motifs. The most common variant (30 strains) differed from the parent peptide at 3 residues (VSIKVGGQIK), but was recognized within 10-fold of the parent peptide. Two variants with conservative changes at anchor residues were both recognized, although the T to A substitution at P2 resulted in a 10-fold reduction in recognition of the variant peptide. All peptides with single changes in non-anchor positions were also recognized, although the P5 variant (G to E) exhibited a decrease in recognition. As the binding was not affected, this probably indicates involvement in T cell recognition.

[00360] Peptides with two changes showed mixed results. In general, peptides with a V substitution at position 3, in combination with another substitution were recognized to the same extent as the corresponding single substitution, indicating the V substitution was tolerated well and is not a TCR contact residue. Combinations including the P2 anchor residue (T to A or N) were not recognized, although the binding of these peptides was also low. Variants with 3 substitutions were generally not recognized well. Two exceptions with very conservative substitutions were noted (Figure 1D). CTL were unable to recognize peptides with four or more substitutions.

[00361] The HLA-A3/11- restricted Env 47 epitope (Figure 1E; VTVYYGVPVWK) was highly conserved, with only 9 variants identified. The most common form observed was the parent peptide (99 strains), while the second most common form, a single anchor substitution observed in 40 strains, was recognized to the same extent as the parent. All the variants were recognized within 10-fold of the parent epitope.

[00362] Taken together, these data show trends towards promiscuous recognition of variant peptides by CTL generated from immunization with a single peptide. In general, changes that disrupted binding also decreased recognition. Recognition was also affected by the position of the change, with potential TCR contact residues (P3-7) exerting a greater effect on recognition than other residues. In general, conservative residue changes were more

widely tolerated than were non-conservative changes. Recognition was also dependent on the number of changes, with progressively lower recognition with a greater number of changes.

[00363] **Recognition after multiple restimulations** The observed recognition of variant peptides by CTL raised against the parent peptide might be due to either promiscuous recognition at the level of a single TCR or simply a mixture of TCRs against the immunizing peptide which are each able to recognize subtly different peptides. To distinguish between these two possibilities, Env 134- or Gag 386-specific T cell lines were generated by stimulating five times with the immunizing peptide, and then tested for recognition of a partial panel of variant peptides. These T cell lines were also characterized for Vβ TCR usage against a panel of antibodies predicted to react with the TCR of the mouse strains utilized for these experiments.

[00364] The data for these peptide-specific lines are shown in Table 5. Because the SU is a measure of the number of cells needed to secrete a defined amount of IFN-γ, a higher SU value would correspond to an enrichment of IFN-γ producing cells. A comparison of one and five peptide stimulations indeed shows an enrichment of CTL specific for the immunizing peptide for both of the peptide lines generated (Table 5A and 5B, first line). The Gag 386 line (Table 5A) also demonstrated increased recognition of all the variant peptides measured except one peptide (ILAEAMSKA) that was never recognized. The Env 134 line also demonstrated enrichment for CTL able to recognize several of the variant peptides (Table 5B).

[00365] To further characterize these lines, we examined them for Vβ usage, utilizing a panel of commercially available antibodies available for mouse TCR Vβ 2-14. To determine background levels for the various TCR Vβ molecules, primary splenocytes from mice that had been immunized with EP HIV-1090 were also examined. The results for the Gag 386 line are shown in Figure 2A. After a single stimulation with the parent peptide, the Gag 386 line showed a mixture of TCR positive populations, including Vβ 3, 5, and 14. After 5 stimulations, those populations had been reduced to background levels, and approximately 50% of the CD8+ cells expressed the Vβ 6 TCR. The Env 134 line showed a similar pattern of multiple TCR positive populations after a single round of stimulation with reduction to background levels after 5 stimulations (data not shown). However, no

single Vβ usage significantly above background could be demonstrated, probably due to lack of the relevant TCR Vβ antibody.

[00366] Both lines were also characterized with regard to the affinity of certain of the variant peptides by titrating the variant peptides examined above (Table 5A and 5B). The data for both the Gag 386 and Env 134 lines are shown in Figure 2B. For the Gag 386 line, the parent peptide along with two single anchor variants (VLAEAMSQI and VLAEAMSQA) showed the highest affinity. Four other peptides demonstrated lower affinity, but still produced IFN-γ in response to higher peptide concentrations. A single peptide (ILAEAMSKA) was not recognized.

[00367] As expected, the parent peptide, which was used to generate the Env 134 line, showed the highest affinity for the TCR. The other 2 variant peptides, KITPLCVTL and QLTPLCVTL, also demonstrated higher affinity, but reduced from the parent peptide by approximately 10-fold and 100-fold, respectively. It was notable that only at the highest peptide concentration examined (1 μg/ml) was any IFN-γ secretion detected for five of the peptides (QITPLCVTL, ELTPLCVTL, KLTPFCVTL, KLTPLCVIL, and KLTPLCVPL). These five peptides showed little or no enrichment of CTL able to recognize them, and exhibited the lowest activity as measured by SU after five restimulations (see Table 5B).

[00368] In summary, these cell lines seem to consist of a narrow, possibly single, TCR population. This TCR population recognizes the parent peptide with the highest affinity, but is also able to recognize a number of other variant peptides with equal or lesser affinity.

[00369] **Recognition of variant peptides by CTL derived from an HIV infected patient.**

[00370] To determine if the same immunological conservation was observed in natural infections, we identified an HIV-infected individual expressing the HLA-A3 allele. The HIV strain and subtype with which this patient was infected is unknown. We had previously shown that T cells from this individual responded to the HLA-A3 restricted epitopes Pol 98 and Env 47. PBL from this patient were examined in an ELISPOT assay to determine if they also showed the capacity for broad cross-reactivity. The data are shown in Figure 3. Although the actual peptide represented in the HIV strain with which this individual is infected is unknown, we observed recognition of a large number of the variant peptides for both Pol 98 (Figure 3A) and Env 47 (Figure 3B). The recognition patterns were remarkably similar for the mouse and patient data (compare Figure 1 and

Figure 3), although the mouse expressed a transgene for HLA-A11 and the patient was HLA-A3.

[00371] **Prediction of Immunological Conservation.** We had observed that the variant peptides that were recognized by CTL raised against the parent epitope had amino acid substitutions that followed previous observations. For example, the anchor residue changes that were tolerated in the variant peptides were also described as anchors that to define the respective HLA supertypes (). In general, conservative substitutions were tolerated at non-anchor residues, while non-conservative substitutions were less well tolerated. These followed closely the prediction model used to identify heteroclitic analogs (Tangri et al).

[00372] Based on these observations, we designed a computer program to predict immunological conservation. For anchor positions, this program utilized the conserved anchor residues described for the A2, A3, and B7 supertypes. For non-anchor positions only conservative substitutions, as defined in Tangri et.al. (), were allowed. All substitutions at non-anchor positions were analyzed independently and all conservative substitutions were allowed regardless of the number of substitutions. Finally, the position of the substitution was not factored into analysis. Each variant was compared with the parent epitope, and its ability to be recognized was predicted as either positive or negative.

[00373] The first sets of epitopes to be evaluated by this program were the five HIV epitopes and variants previously described. For the Env 134 epitope, the program predicted that 13 of the variant peptides should be immunologically conserved, while 6 should not be recognized. Comparison of the observed immunological data with the prediction showed that the program predicted correctly for 14 of the peptides and incorrectly for 5. Of the incorrect predictions, in two cases the program predicted negative results for peptides that were recognized, while in 3 cases the program predicted positive results for peptides that were not recognized. A similar analysis was performed for all five peptides. Of 101 total variant peptides, 68 were correctly identified (67%). The discordant data were fairly evenly split between peptides incorrectly predicted negative (15) and those incorrectly predicted positive (18).

[00374] As noted previously, the more substitutions present in a variant peptide, the lower the likelihood of its immunogenicity. Since the prediction program treated all substitutions independently, and did not take into account the number of substitutions, we hypothesized

that prediction of single substitutions would be more accurate. Indeed, the immunogenicity of 38 of 47 single substitution variants (80%) was correctly predicted.

[00375] With the limitations of the program in mind, it is useful to predict the recognition of the variants for a package of HLA-A2, -A3, and -B7 supertype epitopes. These epitopes had been identified as being well conserved in Clade B variants. When comparing the conservation of this group of epitopes based on sequence identity versus immunological conservation, it is interesting to note that the predicted recognition gains taking into account immunological conservation are significant (Table 6).

[00376] This particular group of 21 epitopes was selected based on their identity conservation in Clade B HIV sequences, with conservation across HIV clades as a secondary consideration. Because of this criteria, the form of epitope chosen as the parent peptide was not the most common variant (e.g. Gag 386, Gag 271, Pol 98). In some cases (e.g., see Gag 386 data), the "parent" epitope and the most common variant were recognized to the same extent. However, in some cases the selection of epitope to include as the "parent" epitope was predicted to make a difference in the immunological conservation. An example of this was the Gag 271 epitope (Figure 4). The variant most commonly seen in clade B sequences was the MTNNPPIPV form, while the most common form of the epitope was MTSNPPIPV. Not all amino acids are considered equal to each other in their ability to substitute (Tangri). For example, asparagine (N) is considered a conservative substitution for serine (S), while the opposite substitution in only considered semi-conserved. When the program calculated immunological conservation using the MTNNPPIPV peptide as the parent peptide, only two variants were predicted to be immunogenic. However, when the immunological conservation was predicted using the MTSNPPIPV peptide, most of the variants were predicted to be recognized (Figure 4). This prediction was tested using HLA-A2 transgenic mice. The results show that if the MTSNPPIPV form of the peptide was utilized in vaccines, approximately 152 of 167 variants would be recognized, while if the MTNNPPIPV form of the epitope was utilized, only 39 of 167 variants would be recognized. This has important implications in epitope selection for vaccine development, and epitope performance can be predicted.

EXAMPLE 3. A PADRE® MOLECULE AS A HELPER EPITOPE FOR ENHANCEMENT OF CTL INDUCTION

[00377] There is increasing evidence that HTL activity is critical for the induction of long lasting CTL responses (Livingston *et al. J. Immunol* 162:3088-3095 (1999); Walter *et al., New Engl. J. Med.* 333:1038-1044 (1995); Hu *et al., J. Exp. Med.* 177:1681-1690 (1993)). Therefore, one or more peptides that bind to HLA class II molecules and stimulate HTLs can be used in accordance with the invention. Accordingly, a preferred embodiment of a vaccine includes a molecule from the PADRE® family of universal T helper cell epitopes (HTL) that target most DR molecules in a manner designed to stimulate helper T cells. For instance, a pan-DR-binding epitope peptide having the formula: aKXVAAZTLKAAa, where "X" is either cyclohexylalanine, phenylalanine, or tyrosine; "Z" is either tryptophan, tyrosine, histidine or asparagine; and "a" is either D-alanine or L-alanine (SEQ ID NO:29), has been found to bind to most HLA-DR alleles, and to stimulate the response of T helper lymphocytes from most individuals, regardless of their HLA type.

[00378] A particularly preferred PADRE® molecule is a synthetic peptide, aKXVAAWTLKAAa (a = D-alanine, X = cyclohexylalanine), containing non-natural amino acids, specifically engineered to maximize both HLA-DR binding capacity and induction of T cell immune responses.

[00379] Alternative preferred PADRE® molecules are the peptides, aKFVAAWTLKAAa, aKYVAAWTLKAAa, aKFVAAYTLKAAa, aKXVAAYTLKAAa, aKYVAAYTLKAAa, aKFVAAHTLKAAa, aKXVAAHTLKAAa, aKYVAAHTLKAAa, aKFVAANTLKAAa, aKXVAANTLKAAa, aKYVAANTLKAAa, AKXVAAWTLKAAA (SEQ ID NO:30), AKFVAAWTLKAAA (SEQ ID NO:31), AKYVAAWTLKAAA (SEQ ID NO:32), AKFVAAYTLKAAA (SEQ ID NO:33), AKXVAAYTLKAAA (SEQ ID NO:34), AKYVAAYTLKAAA (SEQ ID NO:35), AKFVAAHTLKAAA (SEQ ID NO:36), AKXVAAHTLKAAA (SEQ ID NO:37), AKYVAAHTLKAAA (SEQ ID NO:38), AKFVAANTLKAAA (SEQ ID NO:39), AKXVAANTLKAAA (SEQ ID NO:40), AKYVAANTLKAAA (SEQ ID NO:41) (a = D-alanine, X = cyclohexylalanine).

[00380] In a preferred embodiment, the PADRE® peptide is amidated. For example, a particularly preferred amidated embodiment of a PADRE® molecule is conventionally written aKXVAAWTLKAAa-NH_2.

[00381] Competitive inhibition assays with purified HLA-DR molecules demonstrated that the PADRE® molecule aKXVAAWTLKAAa-NH_2 binds with high or intermediate affinity (IC_{50} ≤1,000 nM) to 15 out of 16 of the most prevalent HLA-DR molecules ((Kawashima *et al., Human Immunology* 59:1-14 (1998); Alexander *et al., Immunity* 1:751-761 (1994)). A comparison of the DR binding capacity of PADRE® and tetanus toxoid (TT) peptide 830-843, a "universal" epitope has been published (Panina-Bordignon *et al., Eur. J. Immunology* 19:2237-2242 (1989)). The TT 830-843 peptide bound to only seven of 16 DR molecules tested, while PADRE® bound 15 of 16. At least 1 of the 15 DR molecules that bind PADRE® is predicted to be present in >95% of all humans. Therefore, this PADRE® molecule is anticipated to induce an HTL response in virtually all patients, despite the extensive polymorphism of HLA-DR molecules in the human population.

[00382] PADRE® has been specifically engineered for optimal immunogenicity for human T cells. Representative data from *in vitro* primary immunizations of normal human T cells with TT 830-843 antigen and the PADRE® molecule aKXVAAWTLKAAa-NH_2 are shown in Figure 1. Peripheral blood mononuclear cells (PBMC) from three normal donors were stimulated with the peptides *in vitro*. Following the third round of stimulation, it was observed that PADRE® generated significant primary T cell responses for all three donors as measured in a standard T cell proliferation assay. With the PADRE® peptide, the 10,000 cpm proliferation level was generally reached with 10 to 100 ng/ml of antigen. In contrast, TT 830-843 antigen generated responses for only 2 out of 3 of the individuals tested. Responses approaching the 10,000 cpm range were reached with about 10,000 ng/ml of antigen. In this respect, it was noted that PADRE® was, on a molar basis, about 100-fold more potent than TT 830-843 antigen for activation of T cell responses.

[00383] Early data from a phase I/II investigator-sponsored trial, conducted at the University of Leiden (C.J.M. Melief), support the principle that the PADRE® molecule aKXVAAWTLKAAa, possibly the amidated aKXVAAWTLKAAa -NH_2, is highly immunogenic in humans (Ressing *et al.*, J. Immunother. 23(2):255-66 (2000)). In this trial, a PADRE® molecule was co-emulsified with various human papilloma virus (HPV)-derived CTL epitopes and was injected into patients with recurrent or residual cervical carcinoma. However, because of the late stage of carcinoma with the study patients, it was expected that these patients were immunocompromised. The patients' immunocompromised status was demonstrated by their low frequency of influenza virus-

specific CTL, reduced levels of CD3 expression, and low incidence of proliferative recall responses after *in vitro* stimulation with conventional antigens. Thus, no efficacy was anticipated in the University of Leiden trial, rather the goal of that trial was essentially to evaluate safety. Safety was, in fact, demonstrated. In addition to a favorable safety profile, PADRE® T cell reactivity was detected in four of 12 patients (Figure 2) in spite of the reduced immune competence of these patients.

[00384] Thus, the PADRE® peptide component(s) of the vaccine bind with broad specificity to multiple allelic forms of HLA-DR molecules. Moreover, PADRE® peptide component(s) bind with high affinity ($IC_{50} \leq 1000$ nM), i.e., at a level of affinity correlated with being immunogenic for HLA Class II restricted T cells. The *in vivo* administration of PADRE® peptide(s) stimulates the proliferation of HTL in normal humans as well as patient populations.

[00385] One or more PADRE® peptide(s) may be included in a composition, e.g., a vaccine, comprising one or more peptides, either as an individual peptide(s), fused to one or more variant peptides, or both.

EXAMPLE 4. CTL RECOGNITION OF ENDOGENOUS PROCESSED ANTIGENS AFTER PRIMING

[00386] This example determines that CTL induced by native or analoged peptide epitopes recognize endogenously synthesized, *i.e.*, native antigens.

[00387] Effector cells isolated from transgenic mice that are immunized with peptide epitopes are re-stimulated *in vitro* using peptide-coated stimulator cells. Six days later, effector cells are assayed for cytotoxicity and the cell lines that contain peptide-specific cytotoxic activity are further re-stimulated. An additional six days later, these cell lines are tested for cytotoxic activity on ^{51}Cr labeled Jurkat-A2.1/K^b target cells in the absence or presence of peptide, and also tested on ^{51}Cr labeled target cells bearing the endogenously synthesized antigen, *i.e.* cells that are stably transfected with HIV expression vectors.

[00388] The result will demonstrate that CTL lines obtained from animals primed with peptide epitope recognize endogenously synthesized HIV antigen. The choice of transgenic mouse model to be used for such an analysis depends upon the epitope(s) that is being evaluated. In addition to HLA-A*0201/K^b transgenic mice, several other transgenic

mouse models including mice with human A11, which may also be used to evaluate A3 epitopes, and B7 alleles have been characterized and others (*e.g.*, transgenic mice for HLA-A1 and A24) are being developed. HLA-DR1 and HLA-DR3 mouse models have also been developed, which may be used to evaluate HTL epitopes.

EXAMPLE 5. ACTIVITY OF CTL-HTL CONJUGATED EPITOPES IN TRANSGENIC MICE

[00389] This example illustrates the induction of CTLs and HTLs in transgenic mice by use of a HIV CTL/HTL peptide conjugate whereby the vaccine composition comprises peptides administered to an HIV-infected patient or an individual at risk for HIV. The peptide composition can comprise multiple CTL and/or HTL epitopes. This analysis demonstrates enhanced immunogenicity that can be achieved by inclusion of one or more HTL epitopes in a vaccine composition. Such a peptide composition can comprise an HTL epitope conjugated to a preferred CTL epitope containing, for example, at least one CTL epitope, or an analog of that epitope. The peptides may be lipidated, if desired.

[00390] Immunization procedures: Immunization of transgenic mice is performed as described (Alexander *et al.*, *J. Immunol.* 159:4753-4761, 1997). For example, A2/K^b mice, which are transgenic for the human HLA A2.1 allele and are useful for the assessment of the immunogenicity of HLA-A*0201 motif- or HLA-A2 supermotif-bearing epitopes, are primed subcutaneously (base of the tail) with a 0.1 ml of peptide in Incomplete Freund's Adjuvant, or if the peptide composition is a lipidated CTL/HTL conjugate, in DMSO/saline or if the peptide composition is a polypeptide, in PBS or Incomplete Freund's Adjuvant. Seven days after priming, splenocytes obtained from these animals are restimulated with syngenic irradiated LPS-activated lymphoblasts coated with peptide.

[00391] Cell lines: Target cells for peptide-specific cytotoxicity assays are Jurkat cells transfected with the HLA-A2.1/K^b chimeric gene (*e.g.*, Vitiello *et al.*, *J. Exp. Med.* 173:1007, 1991).

[00392] *In vitro* CTL activation: One week after priming, spleen cells ($30x10^6$ cells/flask) are co-cultured at 37°C with syngeneic, irradiated (3000 rads), peptide coated

lymphoblasts ($10x10^6$ cells/flask) in 10 ml of culture medium/T25 flask. After six days, effector cells are harvested and assayed for cytotoxic activity.

[00393] Assay for cytotoxic activity: Target cells (1.0 to $1.5x10^6$) are incubated at 37°C in the presence of 200 µl of ^{51}Cr. After 60 minutes, cells are washed three times and resuspended in R10 medium. Peptide is added where required at a concentration of 1 µg/ml. For the assay, 10^4 ^{51}Cr-labeled target cells are added to different concentrations of effector cells (final volume of 200 µl) in U-bottom 96-well plates. After a 6 hour incubation period at 37°C, a 0.1 ml aliquot of supernatant is removed from each well and radioactivity is determined in a Micromedic automatic gamma counter. The percent specific lysis is determined by the formula: percent specific release = 100 x (experimental release - spontaneous release)/(maximum release - spontaneous release). To facilitate comparison between separate CTL assays run under the same conditions, % ^{51}Cr release data is expressed as lytic units/10^6 cells. One lytic unit is arbitrarily defined as the number of effector cells required to achieve 30% lysis of 10,000 target cells in a 6 hour ^{51}Cr release assay. To obtain specific lytic units/10^6, the lytic units/10^6 obtained in the absence of peptide is subtracted from the lytic units/10^6 obtained in the presence of peptide. For example, if 30% ^{51}Cr release is obtained at the effector (E): target (T) ratio of 50:1 (i.e., $5x10^5$ effector cells for 10,000 targets) in the absence of peptide and 5:1 (i.e., $5x10^4$ effector cells for 10,000 targets) in the presence of peptide, the specific lytic units would be: $[(1/50,000)-(1/500,000)] \times 10^6 = 18$ LU.

[00394] The results are analyzed to assess the magnitude of the CTL responses of animals injected with the immunogenic CTL/HTL conjugate vaccine preparation and are compared to the magnitude of the CTL response achieved using the CTL epitope as outlined in above. Analyses similar to this may be performed to evaluate the immunogenicity of peptide conjugates containing multiple CTL epitopes and/or multiple HTL epitopes. In accordance with these procedures it is found that a CTL response is induced, and concomitantly that an HTL response is induced upon administration of such compositions.

EXAMPLE 6. SELECTION OF CTL AND HTL EPITOPES FOR INCLUSION IN AN HIV-SPECIFIC VACCINE.

[00395] This example illustrates the procedure for the selection of peptide epitopes for vaccine compositions of the invention. The peptides in the composition can be in the form of a nucleic acid sequence, either single or one or more sequences (*i.e.*, minigene) that encodes peptide(s), or can be single and/or polyepitopic peptides.

[00396] The following principles are utilized when selecting an array of epitopes for inclusion in a vaccine composition. Each of the following principles is balanced in order to make the selection.

[00397] Epitopes are selected which, upon administration, mimic immune responses that correlate with virus clearance. For example, if it has been observed that patients who clear HIV generate an immune response to at least 3 epitopes on at least one HIV antigen, then 3-4 epitopes should be included for HLA class I. A similar rationale is used to determine HLA class II epitopes.

[00398] When selecting an array of HIV epitopes, it is preferred that at least some of the epitopes are derived from early and late proteins. The early proteins of HIV are expressed when the virus is replicating, either following acute or dormant infection. Therefore, it is particularly preferred to use epitopes from early stage proteins to alleviate disease manifestations at the earliest stage possible.

[00399] Epitopes are often selected that have a binding affinity of an IC_{50} of 500 nM or less for an HLA class I molecule, or for class II, an IC_{50} of 1000 nM or less.

[00400] Sufficient supermotif bearing peptides, or a sufficient array of allele-specific motif bearing peptides, are selected to give broad population coverage. For example, epitopes are selected to provide at least 80% population coverage. A Monte Carlo analysis, a statistical evaluation known in the art, can be employed to assess breadth, or redundancy, of population coverage.

[00401] When creating a polyepitopic compositions, *e.g.* a minigene, it is typically desirable to generate the smallest peptide possible that encompasses the epitopes of interest. The principles employed are similar, if not the same, as those employed when selecting a peptide comprising nested epitopes.

[00402] In cases where the sequences of multiple variants of the same target protein are available, potential peptide epitopes can also be selected on the basis of their conservancy.

For example, a criterion for conservancy may define that the entire sequence of an HLA class I binding peptide or the entire 9-mer core of a class II binding peptide be conserved in a designated percentage of the sequences evaluated for a specific protein antigen.

[00403] Peptide epitopes for inclusion in vaccine compositions are, for example, selected from those listed in Tables 6-9 or Figures 1A-4. A vaccine composition comprised of selected peptides, when administered, is safe, efficacious, and elicits an immune response similar in magnitude of an immune response that clears an acute HIV infection.

EXAMPLE 7. CONSTRUCTION OF MINIGENE MULTI-EPITOPE DNA PLASMIDS

[00404] This example provides general guidance for the construction of a minigene expression plasmid. Minigene plasmids may, of course, contain various configurations of CTL and/or HTL epitopes or epitope analogs as described herein. Expression plasmids have been constructed and evaluated as described, for example, in co-pending U.S.S.N. 09/311,784 filed 5/13/99 and in Ishioka *et al.*, *J. Immunol.* 162:3915-3925, 1999. An example of such a plasmid for the expression of HIV epitopes is shown in Figure 2, which illustrates the orientation of HIV peptide epitopes in a minigene construct.

[00405] A minigene expression plasmid typically includes multiple CTL and HTL peptide epitopes. In the present example, HLA-A2, -A3, -B7 supermotif-bearing peptide epitopes and HLA-A1 and -A24 motif-bearing peptide epitopes are used in conjunction with DR supermotif-bearing epitopes and/or DR3 epitopes (Figure 2). Preferred epitopes are identified, for example, in Tables 6-9 and Figures 1A-4. HLA class I supermotif or motif-bearing peptide epitopes derived from multiple HIV antigens, are selected such that multiple supermotifs/motifs are represented to ensure broad population coverage. Similarly, HLA class II epitopes are selected from multiple HIV antigens to provide broad population coverage, *i.e.* both HLA DR-1-4-7 supermotif-bearing epitopes and HLA DR-3 motif-bearing epitopes are selected for inclusion in the minigene construct. The selected CTL and HTL epitopes are then incorporated into a minigene for expression in an expression vector.

[00406] Such a construct may additionally include sequences that direct the HTL epitopes to the endoplasmic reticulum. For example, the Ii protein may be fused to one or more HTL epitopes as described in co-pending application U.S.S.N. 09/311,784 filed 5/13/99,

wherein the CLIP sequence of the Ii protein is removed and replaced with an HLA class II epitope sequence os that HLA class II epitope is directed to the endoplasmic reticulum, where the epitope binds to an HLA class II molecules.

[00407] This example illustrates the methods to be used for construction of a minigene-bearing expression plasmid. Other expression vectors that may be used for minigene compositions are available and known to those of skill in the art.

[00408] The minigene DNA plasmid contains a consensus Kozak sequence and a consensus murine kappa Ig-light chain signal sequence followed by CTL and/or HTL epitopes selected in accordance with principles disclosed herein. The construct can also include, for example, The sequence encodes an open reading frame fused to the Myc and His antibody epitope tag coded for by the pcDNA 3.1 Myc-His vector.

[00409] Overlapping oligonucleotides, for example eight oligonucleotides, averaging approximately 70 nucleotides in length with 15 nucleotide overlaps, are synthesized and HPLC-purified. The oligonucleotides encode the selected peptide epitopes as well as appropriate linker nucleotides, Kozak sequence, and signal sequence. The final multiepitope minigene is assembled by extending the overlapping oligonucleotides in three sets of reactions using PCR. A Perkin/Elmer 9600 PCR machine is used and a total of 30 cycles are performed using the following conditions: 95°C for 15 sec, annealing temperature (5° below the lowest calculated Tm of each primer pair) for 30 sec, and 72°C for 1 min.

[00410] For the first PCR reaction, 5 μg of each of two oligonucleotides are annealed and extended: Oligonucleotides 1+2, 3+4, 5+6, and 7+8 are combined in 100 μl reactions containing *Pfu* polymerase buffer (1x= 10 mM KCL, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-chloride, pH 8.75, 2 mM $MgSO_4$, 0.1% Triton X-100, 100 μg/ml BSA), 0.25 mM each dNTP, and 2.5 U of *Pfu* polymerase. The full-length dimer products are gel-purified, and two reactions containing the product of 1+2 and 3+4, and the product of 5+6 and 7+8 are mixed, annealed, and extended for 10 cycles. Half of the two reactions are then mixed, and 5 cycles of annealing and extension carried out before flanking primers are added to amplify the full length product for 25 additional cycles. The full-length product is gel-purified and cloned into pCR-blunt (Invitrogen) and individual clones are screened by sequencing.

EXAMPLE 8. THE PLASMID CONSTRUCT AND THE DEGREE TO WHICH IT INDUCES IMMUNOGENICITY.

[00411] The degree to which a plasmid construct, for example a plasmid constructed in accordance as above is able to induce immunogenicity can be evaluated *in vitro* by testing for epitope presentation by APC following transduction or transfection of the APC with an epitope-expressing nucleic acid construct. Such a study determines "antigenicity" and allows the use of human APC. The assay determines the ability of the epitope to be presented by the APC in a context that is recognized by a T cell by quantifying the density of epitope-HLA class I complexes on the cell surface. Quantitation can be performed by directly measuring the amount of peptide eluted from the APC (*see, e.g.*, Sijts *et al.*, *J. Immunol.* 156:683-692, 1996; Demotz *et al.*, *Nature* 342:682-684, 1989); or the number of peptide-HLA class I complexes can be estimated by measuring the amount of lysis or lymphokine release induced by infected or transfected target cells, and then determining the concentration of peptide necessary to obtained equivalent levels of lysis or lymphokine release (*see, e.g.*, Kageyama *et al.*, *J. Immunol.* 154:567-576, 1995).

[00412] Alternatively, immunogenicity can be evaluated through *in vivo* injections into mice and subsequent *in vitro* assessment of CTL and HTL activity, which are analysed using cytotoxicity and proliferation assays, respectively, as detailed *e.g.*, in copending U.S.S.N. 09/311,784 filed 5/13/99 and Alexander *et al.*, *Immunity* 1:751-761, 1994.

[00413] For example, to assess the capacity of a DNA minigene construct (*e.g.*, a pMin minigene construct generated as decribed in U.S.S.N. 09/311,784) containing at least one HLA-A2 supermotif peptide to induce CTLs *in vivo*, HLA-A2.1/K^b transgenic mice, for example, are immunized intramuscularly with 100 μg of naked cDNA. As a means of comparing the level of CTLs induced by cDNA immunization, a control group of animals is also immunized with an actual peptide composition that comprises multiple epitopes synthesized as a single polypeptide as they would be encoded by the minigene.

[00414] Splenocytes from immunized animals are stimulated twice with each of the respective compositions (peptide epitopes encoded in the minigene or the polyepitopic peptide), then assayed for peptide-specific cytotoxic activity in a ^{51}Cr release assay. The results indicate the magnitude of the CTL response directed against the A2-restricted epitope, thus indicating the *in vivo* immunogenicity of the minigene vaccine and polyepitopic vaccine. It is, therefore, found that the minigene elicits immune responses

directed toward the HLA-A2 supermotif peptide epitopes as does the polyepitopic peptide vaccine. A similar analysis is also performed using other HLA-A3 and HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 and HLA-B7 motif or supermotif epitopes.

[00415] To assess the capacity of a class II epitope encoding minigene to induce HTLs *in vivo*, DR transgenic mice, or for those epitope that cross react with the appropriate mouse MHC molecule, I-A^b-restricted mice, for example, are immunized intramuscularly with 100 μg of plasmid DNA. As a means of comparing the level of HTLs induced by DNA immunization, a group of control animals is also immunized with an actual peptide composition emulsified in complete Freund's adjuvant. CD4+ T cells, *i.e.* HTLs, are purified from splenocytes of immunized animals and stimulated with each of the respective compositions (peptides encoded in the minigene). The HTL response is measured using a ^{3}H-thymidine incorporation proliferation assay, (*see, e.g.*, Alexander et al. Immunity 1:751-761, 1994). The results indicate the magnitude of the HTL response, thus demonstrating the *in vivo* immunogenicity of the minigene.

[00416] DNA minigenes, constructed as described above or below, may also be evaluated as a vaccine in combination with a boosting agent using a prime boost protocol. The boosting agent can consist of recombinant protein (*e.g.*, Barnett *et al.*, *Aids Res. and Human Retroviruses* 14, Supplement 3:S299-S309, 1998) or recombinant vaccinia, for example, expressing a minigene or DNA encoding the complete protein of interest (*see, e.g.*, Hanke *et al.*, *Vaccine* 16:439-445, 1998; Sedegah *et al.*, *Proc. Natl. Acad. Sci USA* 95:7648-53, 1998; Hanke and McMichael, *Immunol. Letters* 66:177-181, 1999; and Robinson *et al.*, *Nature Med.* 5:526-34, 1999).

[00417] For example, the efficacy of the DNA minigene used in a prime boost protocol is initially evaluated in transgenic mice. In this example, A2.1/K^b transgenic mice are immunized IM with 100 μg of a DNA minigene encoding the immunogenic peptides including at least one HLA-A2 supermotif-bearing peptide. After an incubation period (ranging from 3-9 weeks), the mice are boosted IP with 10^7 pfu/mouse of a recombinant vaccinia virus expressing the same sequence encoded by the DNA minigene. Control mice are immunized with 100 μg of DNA or recombinant vaccinia without the minigene sequence, or with DNA encoding the minigene, but without the vaccinia boost. After an additional incubation period of two weeks, splenocytes from the mice are immediately assayed for peptide-specific activity in an ELISPOT assay. Additionally, splenocytes are

stimulated *in vitro* with the A2-restricted peptide epitopes encoded in the minigene and recombinant vaccinia, then assayed for peptide-specific activity in an IFN-γ ELISA.

[00418] It is found that the minigene utilized in a prime-boost protocol elicits greater immune responses toward the HLA-A2 supermotif peptides than with DNA alone. Such an analysis can also be performed using HLA-A11 or HLA-B7 transgenic mouse models to assess CTL induction by HLA-A3 or HLA-B7 motif or supermotif epitopes.

[00419] The use of prime boost protocols in humans is described in below.

EXAMPLE 9. PEPTIDE COMPOSITION FOR PROPHYLACTIC USES

[00420] Vaccine compositions of the present invention can be used to prevent HIV infection in persons who are at risk for such infection. For example, a polyepitopic peptide epitope composition (or a nucleic acid comprising the same) containing multiple CTL and HTL epitopes, which are also selected to target greater than 80% of the population, is administered to individuals at risk for HIV infection.

[00421] For example, a peptide-based composition can be provided as a single polypeptide that encompasses multiple epitopes. The vaccine is typically administered in a physiological solution that comprises an adjuvant, such as Incomplete Freunds Adjuvant. The dose of peptide for the initial immunization is from about 1 to about 50,000 μg, generally 100-5,000 μg, for a 70 kg patient. The initial administration of vaccine is followed by booster dosages at 4 weeks followed by evaluation of the magnitude of the immune response in the patient, by techniques that determine the presence of epitope-specific CTL populations in a PBMC sample. Additional booster doses are administered as required. The composition is found to be both safe and efficacious as a prophylaxis against HIV infection.

[00422] Alternatively, a composition typically comprising transfecting agents can be used for the administration of a nucleic acid-based vaccine in accordance with methodologies known in the art and disclosed herein.

EXAMPLE 10. POLYEPITOPIC VACCINE COMPOSITIONS DERIVED FROM NATIVE HIV SEQUENCES

[00423] A native HIV polyprotein sequence is screened, preferably using computer algorithms defined for each class I and/or class II supermotif or motif, to identify "relatively short" regions of the polyprotein that comprise multiple epitopes and is preferably less in length than an entire native antigen. This relatively short sequence that contains multiple distinct, even overlapping, epitopes is selected and used to generate a minigene construct. The construct is engineered to express the peptide, which corresponds to the native protein sequence. The "relatively short" peptide is generally less than 250 amino acids in length, often less than 100 amino acids in length, preferably less than 75 amino acids in length, and more preferably less than 50 amino acids in length. The protein sequence of the vaccine composition is selected because it has maximal number of epitopes contained within the sequence, *i.e.*, it has a high concentration of epitopes. As noted herein, epitope motifs may be nested or overlapping, for example, two 9-mer epitopes and one 10-mer epitope can be present in a 10 amino acid peptide. Such a vaccine composition is administered for therapeutic or prophylactic purposes.

[00424] The vaccine composition will preferably include, for example, three CTL epitopes and at least one HTL epitope from HIV. This polyepitopic native sequence is administered either as a peptide or as a nucleic acid sequence which encodes the peptide. Alternatively, an analog can be made of this native sequence, whereby one or more of the epitopes comprise substitutions that alter the cross-reactivity and/or binding affinity properties of the polyepitopic peptide.

[00425] The embodiment of this example provides for the possibility that an as yet undiscovered aspect of immune system processing will apply to the native nested sequence and thereby facilitate the production of therapeutic or prophylactic immune response-inducing vaccine compositions. Additionally such an embodiment provides for the possibility of motif-bearing epitopes for an HLA makeup that is presently unknown. Furthermore, this embodiment (absent analogs) directs the immune response to multiple peptide sequences that are actually present in native HIV antigens thus avoiding the need to evaluate any junctional epitopes. Lastly, the embodiment provides an economy of scale when producing nucleic acid vaccine compositions.

[00426] Related to this embodiment, computer programs can be derived in accordance with principles in the art, which identify in a target sequence, the greatest number of epitopes per sequence length.

EXAMPLE 11. POLYEPITOPIC VACCINE COMPOSITIONS DIRECTED TO MULTIPLE DISEASES

[00427] The HIV peptide epitopes of the present invention are used in conjunction with peptide epitopes from target antigens related to one or more other diseases, to create a vaccine composition that is useful for the prevention or treatment of HIV as well as the one or more other disease(s). Examples of the other diseases include, but are not limited to, HCV and HBV.

[00428] For example, a polyepitopic peptide composition comprising multiple CTL and HTL epitopes that target greater than 98% of the population may be created for administration to individuals at risk for both HBV and HIV infection. The composition can be provided as a single polypeptide that incorporates the multiple epitopes from the various disease-associated sources, or can be administered as a composition comprising one or more discrete epitopes.

EXAMPLE 12. USE OF PEPTIDES TO EVALUATE AN IMMUNE RESPONSE

[00429] Peptides of the invention may be used to analyze an immune response for the presence of specific CTL or HTL populations directed to HIV. Such an analysis may be performed in a manner as that described by Ogg *et al., Science* 279:2103-2106, 1998. In the following example, peptides in accordance with the invention are used as a reagent for diagnostic or prognostic purposes, not as an immunogen.

[00430] In this example highly sensitive human leukocyte antigen tetrameric complexes ("tetramers") are used for a cross-sectional analysis of, for example, HIV HLA-A*0201-specific CTL frequencies from HLA A*0201-positive individuals at different stages of infection or following immunization using an HIV peptide containing an A*0201 motif. Tetrameric complexes are synthesized as described (Musey *et al., N. Engl. J. Med.* 337:1267, 1997). Briefly, purified HLA heavy chain (A*0201 in this example) and β2-

microglobulin are synthesized by means of a prokaryotic expression system. The heavy chain is modified by deletion of the transmembrane-cytosolic tail and COOH-terminal addition of a sequence containing a BirA enzymatic biotinylation site. The heavy chain, β2-microglobulin, and peptide are refolded by dilution. The 45-kD refolded product is isolated by fast protein liquid chromatography and then biotinylated by BirA in the presence of biotin (Sigma, St. Louis, Missouri), adenosine 5'triphosphate and magnesium. Streptavidin-phycoerythrin conjugate is added in a 1:4 molar ratio, and the tetrameric product is concentrated to 1 mg/ml. The resulting product is referred to as tetramer-phycoerythrin.

[00431] For the analysis of patient blood samples, approximately one million PBMCs are centrifuged at 300 x g for 5 minutes and resuspended in 50 μl of cold phosphate-buffered saline. Tri-color analysis is performed with the tetramer-phycoerythrin, along with anti-CD8-Tricolor, and anti-CD38. The PBMCs are incubated with tetramer and antibodies on ice for 30 to 60 min and then washed twice before formaldehyde fixation. Gates are applied to contain >99.98% of control samples. Controls for the tetramers include both A*0201-negative individuals and A*0201-positive uninfected donors. The percentage of cells stained with the tetramer is then determined by flow cytometry. The results indicate the number of cells in the PBMC sample that contain epitope-restricted CTLs, thereby readily indicating the extent of immune response to the HIV epitope, and thus the stage of infection with HIV, the status of exposure to HIV, or exposure to a vaccine that elicits a protective or therapeutic response.

EXAMPLE 13. USE OF PEPTIDE EPITOPES TO EVALUATE RECALL RESPONSES

[00432] The peptide epitopes of the invention are used as reagents to evaluate T cell responses, such as acute or recall responses, in patients. Such an analysis may be performed on patients who have recovered from infection, who are chronically infected with HIV, or who have been vaccinated with an HIV vaccine.

[00433] For example, the class I restricted CTL response of persons who have been vaccinated may be analyzed. The vaccine may be any HIV vaccine. PBMC are collected from vaccinated individuals and HLA typed. Appropriate peptide epitopes of the invention that, optimally, bear supermotifs to provide cross-reactivity with multiple HLA

supertype family members, are then used for analysis of samples derived from individuals who bear that HLA type.

[00434] PBMC from vaccinated individuals are separated on Ficoll-Histopaque density gradients (Sigma Chemical Co., St. Louis, MO), washed three times in HBSS (GIBCO Laboratories), resuspended in RPMI-1640 (GIBCO Laboratories) supplemented with L-glutamine (2mM), penicillin (50U/ml), streptomycin (50 μg/ml), and Hepes (10mM) containing 10% heat-inactivated human AB serum (complete RPMI) and plated using microculture formats. A synthetic peptide comprising an epitope of the invention is added at 10 μg/ml to each well and HBV core 128-140 epitope is added at 1 μg/ml to each well as a source of T cell help during the first week of stimulation.

[00435] In the microculture format, 4 x 10^5 PBMC are stimulated with peptide in 8 replicate cultures in 96-well round bottom plate in 100 μl/well of complete RPMI. On days 3 and 10, 100 ml of complete RPMI and 20 U/ml final concentration of rIL-2 are added to each well. On day 7 the cultures are transferred into a 96-well flat-bottom plate and restimulated with peptide, rIL-2 and 10^5 irradiated (3,000 rad) autologous feeder cells. The cultures are tested for cytotoxic activity on day 14. A positive CTL response requires two or more of the eight replicate cultures to display greater than 10% specific ^{51}Cr release, based on comparison with uninfected control subjects as previously described (Rehermann, *et al., Nature Med.* 2:1104,1108, 1996; Rehermann *et al., J. Clin. Invest.* 97:1655-1665, 1996; and Rehermann *et al. J. Clin. Invest.* 98:1432-1440, 1996).

[00436] Target cell lines are autologous and allogeneic EBV-transformed B-LCL that are either purchased from the American Society for Histocompatibility and Immunogenetics (ASHI, Boston, MA) or established from the pool of patients as described (Guilhot, *et al. J. Virol. 66*:2670-2678, 1992).

[00437] Cytotoxicity assays are performed in the following manner. Target cells consist of either allogeneic HLA-matched or autologous EBV-transformed B lymphoblastoid cell line that are incubated overnight with the synthetic peptide epitope of the invention at 10 μM, and labeled with 100 μCi of ^{51}Cr (Amersham Corp., Arlington Heights, IL) for 1 hour after which they are washed four times with HBSS.

[00438] Cytolytic activity is determined in a standard 4-h, split well ^{51}Cr release assay using U-bottomed 96 well plates containing 3,000 targets/well. Stimulated PBMC are tested at effector/target (E/T) ratios of 20-50:1 on day 14. Percent cytotoxicity is determined from the formula: 100 x [(experimental release-spontaneous

release)/maximum release-spontaneous release)]. Maximum release is determined by lysis of targets by detergent (2% Triton X-100; Sigma Chemical Co., St. Louis, MO). Spontaneous release is <25% of maximum release for all experiments.

[00439] The results of such an analysis indicate the extent to which HLA-restricted CTL populations have been stimulated by previous exposure to HIV or an HIV vaccine.

[00440] The class II restricted HTL responses may also be analyzed. Purified PBMC are cultured in a 96-well flat bottom plate at a density of $1.5x10^5$ cells/well and are stimulated with 10 μg/ml synthetic peptide, whole antigen, or PHA. Cells are routinely plated in replicates of 4-6 wells for each condition. After seven days of culture, the medium is removed and replaced with fresh medium containing 10U/ml IL-2. Two days later, 1 μCi ^{3}H-thymidine is added to each well and incubation is continued for an additional 18 hours. Cellular DNA is then harvested on glass fiber mats and analyzed for ^{3}H-thymidine incorporation. Antigen-specific T cell proliferation is calculated as the ratio of ^{3}H-thymidine incorporation in the presence of antigen divided by the ^{3}H-thymidine incorporation in the absence of antigen.

EXAMPLE 14. INDUCTION OF SPECIFIC CTL RESPONSE IN HUMANS

[00441] A human clinical trial for an immunogenic composition comprising CTL and HTL epitopes of the invention is set up as an IND Phase I, dose escalation study and carried out as a randomized, double-blind, placebo-controlled trial. Such a trial is designed, for example, as follows:

[00442] A total of about 27 subjects are enrolled and divided into 3 groups:

Group I: 3 subjects are injected with placebo and 6 subjects are injected with 5 μg of peptide composition;

Group II: 3 subjects are injected with placebo and 6 subjects are injected with 50 μg peptide composition;

Group III: 3 subjects are injected with placebo and 6 subjects are injected with 500 μg of peptide composition.

[00443] After 4 weeks following the first injection, all subjects receive a booster inoculation at the same dosage.

[00444] The endpoints measured in this study relate to the safety and tolerability of the peptide composition as well as its immunogenicity. Cellular immune responses to the peptide composition are an index of the intrinsic activity of this the peptide composition, and can therefore be viewed as a measure of biological efficacy. The following summarize the clinical and laboratory data that relate to safety and efficacy endpoints.

[00445] Safety: The incidence of adverse events is monitored in the placebo and drug treatment group and assessed in terms of degree and reversibility.

[00446] Evaluation of Vaccine Efficacy: For evaluation of vaccine efficacy, subjects are bled before and after injection. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

[00447] The vaccine is found to be both safe and efficacious.

EXAMPLE 15. PHASE II TRIALS IN PATIENTS INFECTED WITH HIV

[00448] Phase II trials are performed to study the effect of administering the CTL-HTL peptide compositions to HIV-infected patients. The main objectives of the trials are to determine an effective dose and regimen for inducing CTLs in chronically infected HIV patients, to establish the safety of inducing a CTL and HTL response in these patients, and to see to what extent activation of CTLs improves the clinical picture of chronically infected HIV patients, as manifested by a reduction in viral load and an increase in $CD4^+$ cells counts. Such a study is designed, for example, as follows:

[00449] The studies are performed in multiple centers. The trial design is an open-label, uncontrolled, dose escalation protocol wherein the peptide composition is administered as a single dose followed six weeks later by a single booster shot of the same dose. The dosages are 50, 500 and 5,000 micrograms per injection. Drug-associated adverse effects (severity and reversibility) are recorded.

[00450] There are three patient groupings. The first group is injected with 50 micrograms of the peptide composition and the second and third groups with 500 and 5,000 micrograms of peptide composition, respectively. The patients within each group range in age from 21-65, include both males and females, and represent diverse ethnic

backgrounds. All of them are infected with HIV for over five years and are HCV, HBV and delta hepatitis virus (HDV) negative, but have positive levels of HIV antigen.

[00451] The viral load and $CD4^+$ levels are monitored to assess the effects of administering the peptide compositions. The vaccine composition is found to be both safe and efficacious in the treatment of HIV infection.

EXAMPLE 16. INDUCTION OF CTL RESPONSES USING A PRIME BOOST PROTOCOL

[00452] A prime boost protocol can also be used for the administration of the vaccine to humans. Such a vaccine regimen can include an initial administration of, for example, naked DNA followed by a boost using recombinant virus encoding the vaccine, or recombinant protein/polypeptide or a peptide mixture administered in an adjuvant.

[00453] For example, the initial immunization is performed using an expression vector, such as that constructed above, in the form of naked nucleic acid administered IM (or SC or ID) in the amounts of 0.5-5 mg at multiple sites. The nucleic acid (0.1 to 1000 μg) can also be administered using a gene gun. Following an incubation period of 3-4 weeks, a booster dose is then administered. The booster is, for example, recombinant fowlpox virus administered at a dose of $5\text{-}10^7$ to $5\text{x}10^9$ pfu. An alternative recombinant virus, such as an MVA, canarypox, adenovirus, or adeno-associated virus, can also be used for the booster, or the polyepitopic protein or a mixture of the peptides can be administered. For evaluation of vaccine efficacy, patient blood samples are obtained before immunization as well as at intervals following administration of the initial vaccine and booster doses of the vaccine. Peripheral blood mononuclear cells are isolated from fresh heparinized blood by Ficoll-Hypaque density gradient centrifugation, aliquoted in freezing media and stored frozen. Samples are assayed for CTL and HTL activity.

[00454] Analysis of the results indicates that a magnitude of sufficient response to achieve protective immunity against HIV is generated.

EXAMPLE 17. ADMINISTRATION OF VACCINE COMPOSITIONS USING DENDRITIC CELLS

[00455] Vaccines comprising peptide epitopes of the invention can be administered using APCs, or "professional" APCs such as DC. In this example, the peptide-pulsed DC are administered to a patient to stimulate a CTL response *in vivo*. In this method, dendritic cells are isolated, expanded, and pulsed with a vaccine comprising peptide CTL and HTL epitopes of the invention. The dendritic cells are infused back into the patient to elicit CTL and HTL responses *in vivo*. The induced CTL and HTL then destroy or facilitate destruction of the specific target cells that bear the proteins from which the epitopes in the vaccine are derived.

[00456] For example, a cocktail of epitope-bearing peptides is administered *ex vivo* to PBMC, or isolated DC therefrom. A pharmaceutical to facilitate harvesting of DC can be used, such as Progenipoietin™ (Monsanto, St. Louis, MO) or GM-CSF/IL-4. After pulsing the DC with peptides and prior to reinfusion into patients, the DC are washed to remove unbound peptides.

[00457] As appreciated clinically, and readily determined by one of skill based on clinical outcomes, the number of DC reinfused into the patient can vary (*see, e.g., Nature Med.* 4:328, 1998; *Nature Med.* 2:52, 1996 and *Prostate* 32:272, 1997). Although 2-50 x 10^6 DC per patient are typically administered, larger number of DC, such as 10^7 or 10^8 can also be provided. Such cell populations typically contain between 50-90% DC.

[00458] In some embodiments, peptide-loaded PBMC are injected into patients without purification of the DC. For example, PBMC containing DC generated after treatment with an agent such as Progenipoietin™ are injected into patients without purification of the DC. The total number of PBMC that are administered often ranges from 10^8 to 10^{10}. Generally, the cell doses injected into patients is based on the percentage of DC in the blood of each patient, as determined, for example, by immunofluorescence analysis with specific anti-DC antibodies. Thus, for example, if Progenipoietin™ mobilizes 2% DC in the peripheral blood of a given patient, and that patient is to receive 5 x 10^6 DC, then the patient will be injected with a total of 2.5 x 10^8 peptide-loaded PBMC. The percent DC mobilized by an agent such as Progenipoietin™ is typically estimated to be between 2-10%, but can vary as appreciated by one of skill in the art.

Ex vivo activation of CTL/HTL responses

[00459] Alternatively, *ex vivo* CTL or HTL responses to HIV antigens can be induced by incubating in tissue culture the patient's, or genetically compatible, CTL or HTL precursor cells together with a source of APC, such as DC, and the appropriate immunogenic peptides. After an appropriate incubation time (typically about 7-28 days), in which the precursor cells are activated and expanded into effector cells, the cells are infused back into the patient, where they will destroy or facilitate destruction of their specific target cells.

[00460] It is understood that the examples and embodiments described herein are for illustrative purposes only and that various modifications or changes in light thereof will be suggested to persons skilled in the art and are to be included within the spirit and purview of this application and scope of the appended claims. All publications, patents, patent applications and sequence listings cited herein are hereby incorporated by reference in their entirety for all purposes.

TABLE 1

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	**T, I,** *L, V, M, S*		**F, W, Y**
A2	**L, I, V, M,** *A, T, Q*		**I, V,** *M, A, T, L*
A3	**V, S, M, A,** *T, L, I*		**R,K**
A24	**Y, F,** *W, I, V, L, M, T*		**F, I,** *Y, W, L, M*
B7	**P**		**V, I, L, F,** *M, W, Y, A*
B27	**R, H, K**		**F, Y, L,** *W, M, I, V, A*
B44	**E,** *D*		**F, W, L, I, M, V, A**
B58	**A, T, S**		**F, W, Y,** *L, I, V, M, A*
B62	**Q, L,** *I, V, M, P*		**F, W, Y,** *M, I, V, L, A*
MOTIFS			
A1	**T, S, M**		**Y**
A1		**D, E,** *A, S*	**Y**
A2.1	**L, M,** *V, Q, I, A, T*		**V,** *L, I, M, A, T*
A3	**L, M, V, I, S, A, T, F,** *C, G, D*		**K, Y, R,** *H, F, A*
A11	**V, T, M, L, I, S, A, G, N,** *C, D, F*		**K,** *R, Y, H*
A24	**Y, F, W,** *M*		**F, L, I, W**
A*3101	**M, V, T,** *A ,L, I, S*		**R,** *K*
A*3301	**M, V, A, L, F,** *I, S, T*		**R, K**
A*6801	**A, V, T,** *M, S, L, I*		**R, K**
B*0702	**P**		**L, M, F,** *W, Y, A, I, V*
B*3501	**P**		**L, M, F, W, Y,** *I, V, A*
B51	**P**		**L, I, V, F,** *W, Y, A, M*
B*5301	**P**		**I, M, F, W, Y,** *A, L, V*
B*5401	**P**		**A, T, I, V,** *L, M, F, W, Y*

Bolded residues are preferred, italicized residues are tolerated: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE 2

SUPERMOTIFS	POSITION	POSITION	POSITION
	2 (Primary Anchor)	3 (Primary Anchor)	C Terminus (Primary Anchor)
A1	**T, I,** *L, V, M, S*		**F, W, Y**
A2	**V,** *Q, A, T*		**I, V,** *L, M, A, T*
A3	**V, S, M, A,** *T, L, I*		**R, K**
A24	**Y, F,** *W, I, V, L, M, T*		**F, I,** *Y, W, L, M*
B7	**P**		**V, I, L, F,** *M, W, Y, A*
B27	**R, H, K**		**F, Y, L,** *W, M, I, V, A*
B58	**A, T, S**		**F, W, Y,** *L, I, V, M, A*
B62	**Q, L,** *I, V, M, P*		**F, W, Y,** *M, I, V, L, A*
MOTIFS			
A1	**T, S, M**		**Y**
A1		**D, E,** *A, S*	**Y**
A2.1	*V, Q, A, T**		**V,** *L, I, M, A, T*
A3.2	**L, M, V, I, S, A, T, F,** *C, G, D*		**K, Y, R,** *H, F, A*
A11	**V, T, M, L, I, S, A, G, N,** *C, D, F*		**K,** *R, H, Y*
A24	**Y ,F, W**		**F, L, I, W**

*If 2 is V, or Q, the C-term is not L

Bolded residues are preferred, italicized residues are tolerated: A peptide is considered motif-bearing if it has primary anchors at each primary anchor position for a motif or supermotif as specified in the above table.

TABLE 3

		POSITION 1	2	3	4	5	6	7	8	C-terminus
SUPERMOTIFS										
A1			1° Anchor T,I,*L,V,M,S*							1° Anchor F,W,Y
A2			1° Anchor L,I,V,M,*A, T,Q*							1° Anchor L,I,V,M,A,T
A3	preferred		1° Anchor V,S,M,A,*T, L,I*	Y,F,W, (4/5)			Y,F,W, (3/5)	Y,F,W, (4/5)	P, (4/5)	1°Anchor R,K
	deleterious	D,E (3/5); P, (5/5)		D,E, (4/5)						
A24			1° Anchor Y,F,*W,I,V, L,M,T*							1° Anchor F,I,*Y,W,L,M*
B7	preferred	F,W,Y (5/5) L,I,V,M, (3/5)	1°Anchor P	F,W,Y (4/5)					F,W,Y, (3/5)	1°Anchor V,I,L,F,*M,W,Y,A*
	deleterious	D,E (3/5); P(5/5); G(4/5); A(3/5); Q,N, (3/5)				D,E, (3/5)	G, (4/5)	Q,N, (4/5)	D,E, (4/5)	
B27			1° Anchor R,H,K							1° Anchor F,Y,L,*W,M,V,A*
B44			1° Anchor E,*D*							1° Anchor F,W,Y,L,I,M,V,A
			1° Anchor							1° Anchor

TABLE 3 (Continued)

		POSITION									
		1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
B58			A,T,S							F,W,Y,*L,I,V,M,A*	
B62			1° Anchor Q,L,*I,V,M,P*							1° Anchor F,W,Y,*M,I,V,L,A*	
MOTIFS											
A1 9-mer	preferred	G,F,Y,W,	1°Anchor S,T,M,	D,E,A,	Y,F,W,		P,	D,E,Q,N,	Y,F,W,	1°Anchor Y	
	deleterious	D,E,		R,H,K,L,I,V M,P,	A,	G,	A,				
A1 9-mer	preferred	G,R,H,K	A,S,T,C,L,I V,M,	1°Anchor D,E,*A,S*	G,S,T,C,		A,S,T,C,	L,I,V,M,	D,E,	1°Anchor Y	
	deleterious	A	R,H,K,D,E, P,Y,F,W,		D,E,	P,Q,N,	R,H,K,	P,G,	G,P,		
A1 10-mer	peferred	Y,F,W,	1°Anchor S,T,M	D,E,A,Q,N,	A,	Y,F,W,Q,N,		P,A,S,T,C,	G,D,E,	P,	1°Anchor Y
	deleterious	G,P,		R,H,K,G,L,I V,M,	D,E,	R,H,K,	Q,N,A	R,H,K,Y,F, W,	R,H,K,	A	

TABLE 3 (Continued)

		POSITION									
		1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
A1 10-mer	preferred	Y,F,W,	S,T,C,L,I,V M,	1°Anchor D,E,*A,S*	A,	Y,F,W,		P,G,	G,	Y,F,W,	1°Anchor Y
	deleterious	R,H,K,	R,H,K,D,E, P,Y,F,W,			P,	G,		P,R,H,K,	Q,N,	
A2.1 9-mer	preferred	Y,F,W,	1°Anchor L,M,*I,V,Q, A,T*	Y,F,W,	S,T,C,	Y,F,W,		A,	P	1°Anchor V,*L,I,M,A,T*	
	deleterious	D,E,P,		D,E,R,K,H			R,K,H	D,E,R,K,H			
A2.1 10-mer	preferred	A,Y,F,W,	1°Anchor L,M,*I,V,Q, A,T*	L,V,I,M,	G,		G,		F,Y,W,L, V,I,M,		1°Anchor V,*L,I,M,A,T*
	deleterious	D,E,P,		D,E,	R,K,H,A,	P,		R,K,H,	D,E,R,K, H,	R,K,H,	

TABLE 3 (Continued)

		POSITION								
		1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
A3	preferred	R,H,K,	1°Anchor L,M,V,I,S, A,T,F,*C,G D*	Y,F,W,	P,R,H,K,Y, F,W,	A,	Y,F,W,		P,	1°Anchor K,Y,R,*H,F,A*	
	deleterious	D,E,P,		D,E							
A11	preferred	A,	1°Anchor V,T,L,M,I, S,A,G,N,*C, D,F*	Y,F,W,	Y,FW,	A,	Y,F,W,	Y,FW,	P,	1°Anchor K,,*RY,H*	
	deleterious	D,E,P,						A	G,		
A24 9-mer	preferred	Y,F,W,R,H,K,	1°Anchor Y,F,W,*M*		S,T,C			Y,F,W,	Y,F,W,	1°Anchor F,L,I,W	
	deleterious	D,E,G,		D,E,	G,	Q,N,P,	D,E,R,H,K,	G,	A,Q,N,		
A24 10-mer	preferred		1°Anchor Y,F,W,*M*		P,	Y,F,W,P,		P,			1°Anchor F,L,I,W
	deleterious			G,D,E	Q,N	R,H,K	D,E	A	Q,N,	D,E,A,	
A3101	preferred	R,H,K,	1°Anchor M,V,T,*A,L, I,S*	Y,F,W,	P,		Y,F,W,	Y,F,W,	A,P,	1°Anchor R,*K*	

TABLE 4 (Continued)

TABLE 3 (Continued)

		POSITION								
		1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
	deleterious	D,E,P,		D,E,		A,D,E,	D,E,	D,E,	D,E,		
A3301	preferred		1°Anchor M,V,A,L,F, *I,S,T*	Y,F,W				A,Y,F,W		1°Anchor R,K	
	deleterious	G,P		D,E							
A6801	preferred	Y,F,W,S,T,C,	1°Anchor A,V,T,*M,S, L,I*			Y,F,W,L,I, V,M		Y,F,W,	P,	1°Anchor R,K	
	deleterious	G,P,		D,E,G,		R,H,K,			A,		
B0702	preferred	R,H,K,F,W,Y,	1°Anchor P	R,H,K,		R,H,K,	R,H,K,	R,H,K,	P,A,	1°Anchor L,M,F,*W,Y,A, I,V*	
	deleterious	D,E,Q,N,P,		D,E,P,	D,E,	D,E,	G,D,E,	Q,N,	D,E,		
B3501	preferred	F,W,Y,L,I,V,M,	1°Anchor P	F,W,Y,				F,W,Y,		1°Anchor L,M,F,W,Y,*I, V,A*	
	deleterious	A,G,P,				G,	G,				

TABLE 3 (Continued)

		POSITION									
		1	2	3	4	5	6	7	8	9 or C-terminus	C-terminus
B51	preferred	L,I,V,M,F,W,Y,	1°Anchor P	F,W,Y,	S,T,C,	F,W,Y,		G,	F,W,Y,	1°Anchor L,I,V,F,*W, Y,A,M*	
	deleterious	A,G,P,D,E,R,H,K, S,T,C,				D,E,	G,	D,E,Q,N,	G,D,E,		
B5301	preferred	L,I,V,M,F,W,Y,	1°Anchor P	F,W,Y,	S,T,C,	F,W,Y,		L,I,V,M,F, W,Y,	F,W,Y,	1°Anchor I,M,F,W,Y, *A,L,V*	
	deleterious	A,G,P,Q,N,					G,	R,H,K,Q,N,	D,E,		
B5401	preferred	F,W,Y,	1°Anchor P	F,W,Y,L,I,V M,		L,I,V,M,		A,L,I,V,M,	F,W,Y,A,P,	1°Anchor A,T,I,V,*L, M,F,W,Y*	
	deleterious	G,P,Q,N,D,E,		G,D,E,S,T,C,		R,H,K,D,E,	D,E,	Q,N,D,G,E,	D,E,		

Italicized residues indicate "tolerated" residues.
The information in Table II is specific for 9-mers unless otherwise specified.
Secondary anchor specificities are designated for each position independently.

Table 4

HLA-supertype	Allele-specific HLA-supertype members	
	Verified[a]	Predicted[b]
A1	A*0101, A*2501, A*2601, A*2602, A*3201	A*0102, A*2604, A*3601, A*4301, A*8001
A2	A*0201, A*0202, A*0203, A*0204, A*0205, A*0206, A*0207, A*0209, A*0214, A*6802, A*6901	A*0208, A*0210, A*0211, A*0212, A*0213
A3	A*0301, A*1101, A*3101, A*3301, A*6801	A*0302, A*1102, A*2603, A*3302, A*3303, A*3401, A*3402, A*6601, A*6602, A*7401
A24	A*2301, A*2402, A*3001	A*2403, A*2404, A*3002, A*3003
B7	B*0702, B*0703, B*0704, B*0705, B*1508, B*3501, B*3502, B*3503, B*3503, B*3504, B*3505, B*3506, B*3507, B*3508, B*5101, B*5102, B*5103, B*5104, B*5105, B*5301, B*5401, B*5501, B*5502, B*5601, B*5602, B*6701, B*7801	B*1511, B*4201, B*5901
B27	B*1401, B*1402, B*1509, B*2702, B*2703, B*2704, B*2705, B*2706, B*3801, B*3901, B*3902, B*7301	B*2701, B*2707, B*2708, B*3802, B*3903, B*3904, B*3905, B*4801, B*4802, B*1510, B*1518, B*1503
B44	B*1801, B*1802, B*3701, B*4402, B*4403, B*4404, B*4001, B*4002, B*4006	B*4101, B*4501, B*4701, B*4901, B*5001
B58	B*5701, B*5702, B*5801, B*5802, B*1516, B*1517	
B62	B*1501, B*1502, B*1513, B*5201	B*1301, B*1302, B*1504, B*1505, B*1506, B*1507, B*1515, B*1520, B*1521, B*1512, B*1514, B*1510

a. Verified alleles include alleles whose specificity has been determined by pool sequencing analysis, peptide binding assays, or by analysis of the sequences of CTL epitopes.

b. Predicted alleles are alleles whose specificity is predicted on the basis of B and F pocket structure to overlap with the supertype specificity.

Table 5. Compiled rankings and similarity assignments.

A		C		D		E		F		G		H		I		K		L	
A	1.0	C	1.0	D	1.0	E	1.0	F	1.0	G	1.0	H	1.0	I	1.0	K	1.0	L	1.0
S	4.5	V	5.5	N	3.5	Q	3.3	L	3.8	S	2.8	Q	2.0	L	2.5	R	2.7	I	3.3
T	4.8	T	6.5	E	4.0	N	4.2	Y	4.3	A	4.2	E	5.8	M	4.5	Q	6.0	M	4.0
P	5.3	A	6.7	Q	6.2	D	4.7	I	4.8	T	4.7	N	6.2	F	5.2	H	6.8	F	4.7
G	5.7	S	7.3	T	7.2	H	5.3	M	6.2	D	6.3	R	8.2	V	5.2	N	7.2	V	5.5
C	9.3	P	8.0	S	7.7	T	8.5	V	6.7	P	7.0	K	8.7	Y	8.8	E	7.5	Y	8.2
V	9.5	I	9.0	H	7.8	K	8.7	W	7.2	N	7.3	P	9.0	T	10.3	D	8.5	H	9.7
D	10.0	Y	10.2	G	8.8	P	9.0	H	10.2	E	8.3	D	9.2	H	11.3	M	9.2	Q	10.5
M	10.0	N	10.7	P	9.2	R	10.5	C	11.2	Q	9.3	S	10.5	A	11.5	T	10.5	W	11.0
N	10.2	F	11.0	A	9.3	S	10.5	T	11.5	H	9.8	T	10.5	C	11.7	S	11.3	I	11.3
E	11.3	G	11.0	K	9.5	V	11.0	R	11.8	C	11.5	Y	11.0	K	12.0	P	11.5	A	11.8
Q	11.3	M	11.0	R	11.5	G	12.0	A	13.0	V	12.0	M	11.2	P	12.0	I	12.2	K	12.2
H	12.3	H	11.3	C	12.0	A	12.2	K	13.3	M	13.5	V	12.2	Q	12.2	L	12.7	P	12.2
Y	13.0	D	12.0	V	13.2	M	12.3	P	13.3	K	13.8	L	13.3	R	13.0	Y	12.8	R	13.2
I	13.7	L	13.2	M	14.3	I	15.0	S	13.8	Y	14.2	I	13.7	S	13.2	V	13.7	C	13.7
L	14.3	E	13.7	Y	15.8	C	15.2	Q	14.0	R	15.5	A	13.8	E	14.0	A	14.3	E	14.2
W	14.3	Q	13.7	I	16.0	L	15.3	N	14.2	I	16.3	G	13.8	N	14.3	G	14.3	N	14.5
F	15.8	W	15.0	L	16.7	Y	15.7	E	15.3	L	16.8	F	15.3	W	14.7	F	14.8	S	15.3
K	15.8	R	16.3	W	18.0	F	17.8	G	16.3	W	17.3	C	17.2	D	15.7	W	15.5	G	16.8
R	17.7	K	17.0	F	18.3	W	18.0	D	18.0	F	18.2	W	17.5	G	17.0	C	17.5	D	17.2

Conserved (1-7) Semi-conserved (7.1-13) Non-conserved (13.1-20)

Table 5 (continued)

M		N		P		Q		R		S		T		V		W		Y	
M	1.0	N	1.0	P	1.0	Q	1.0	R	1.0	S	1.0	T	1.0	V	1.0	W	1.0	Y	1.0
L	3.8	D	3.2	T	3.5	E	3.2	K	2.0	T	4.0	P	4.7	L	5.3	F	4.2	F	6.2
I	5.0	E	4.5	S	6.0	H	3.3	H	5.8	G	4.2	S	5.0	I	5.5	Y	4.5	W	6.3
V	5.2	Q	5.8	A	6.3	N	6.0	Q	6.7	A	4.7	A	5.7	M	5.5	L	6.3	H	7.3
F	7.0	H	6.8	H	7.2	K	7.0	E	8.7	P	5.3	N	6.5	P	7.8	M	6.5	M	8.0
K	9.7	T	7.2	Q	7.2	D	7.2	N	8.7	N	6.5	D	7.8	T	8.2	R	8.7	L	9.7
Q	10.3	S	7.7	N	7.8	R	8.3	M	9.2	D	8.0	E	8.3	F	8.5	I	9.2	T	10.0
R	10.5	K	8.5	D	9.2	P	8.5	D	10.0	C	9.3	G	8.8	A	10.0	H	10.2	I	10.3
Y	10.8	P	8.7	E	9.5	M	10.5	P	10.2	E	9.3	H	10.5	C	10.7	V	10.3	C	10.7
H	11.0	R	10.5	G	10.2	T	10.8	W	10.3	H	10.2	Q	10.5	H	11.0	K	10.5	V	11.0
A	11.2	A	10.7	V	11.2	V	11.3	S	11.3	Q	10.5	V	10.5	Q	11.0	Q	12.3	P	11.5
T	11.3	G	11.8	C	11.8	S	12.3	T	12.0	K	12.3	C	12.2	E	11.3	S	12.3	A	11.7
P	11.8	V	12.8	M	12.2	A	13.0	I	12.5	Y	13.5	K	12.2	N	12.3	A	12.5	N	11.7
W	12.7	C	13.3	R	13.2	G	13.3	L	13.0	R	13.7	M	12.3	S	12.7	P	12.7	Q	12.0
C	13.0	M	13.8	K	13.5	L	13.8	Y	13.0	M	14.5	I	13.7	G	13.2	T	12.7	S	12.2
E	13.2	Y	15.2	Y	13.8	I	14.2	V	13.7	Y	14.7	Y	14.3	Y	14.0	N	13.2	K	12.7
N	14.2	I	16.0	I	16.0	Y	15.0	F	14.3	I	16.0	R	15.5	D	14.8	C	13.3	R	13.0
S	15.0	L	16.5	L	16.3	C	16.3	A	15.2	W	16.5	W	16.3	K	15.2	E	15.7	E	14.2
G	16.5	W	17.7	W	16.3	F	17.3	G	15.2	L	17.5	L	16.5	W	15.7	G	16.7	G	14.7
D	16.8	F	18.9	F	17.8	W	17.5	C	17.3	F	18.3	F	17.7	R	16.3	D	17.9	D	16.0

Conserved (1-7) Semi-conserved (7.1-13) Non-conserved (13.1-20)

Table 6. Recognition of variant peptides by CTL generated after one and five stimulations with the parent peptide.

A. Gag 386 (VLAEAMSQV)

Peptide Sequence	Binding IC50 (nM)	1 Stimulation (SU)	5 Stimulations (SU)
VLAEAMSQV	49.9	31.6	222.0
VLAEAMSQA	23.8	17.0	133.5
VLAEAMSQI	70.9	21.2	246.1
VLAEAMSKV	230.5	10.8	130.9
VLAEAMSKA	69.4	NT	36.6
ILAEAMSQA	29.3	4.0	49.7
ILAEAMSKA	72.4	--	--
VLAEAMAAA	17	16.3	90.3

B. Env 134 (KLTPLCVTL)

KLTPLCVTL	77.0	278.4	683.6
KITPLCVTL	461	231.8	700.8
QLTPLCVTL	63.6	166.2	361.5
QITPLCVTL	975	105.0	166.9
ELTPLCVTL	7190	91.7	100.0
KLTPFCVTL	87.3	36.1	75.4
KLTPLCVIL	356	77.2	29.1
KLTPLCVPL	14.6	9.6	14.8

Table 7. Conservation of EP HIV-1090 epitopes across clades, calculated as identity or immunological conservation

		Total		Clade B		Clade C	
Protein	Sequence	Identity	Imm. Cons.	Identity	Imm. Cons.	Identity	Imm. Cons.
Pol 498	ILKEPVHGV	62%	87%	77%	86%	74%	95%
Gag 386	VLAEAMSQV	32%	93%	68%	91%	5%	94%
Pol 448	KLVGKLNWA	95%	96%	95%	95%	95%	98%
Env 134	KLTPLCVTL	80%	93%	90%	95%	89%	98%
Vpr 62	RILQQLLFI	51%	93%	68%	91%	61%	95%
Nef 221	LTFGWCFKL	49%	74%	77%	91%	47%	81%
Gag 271	MTNNPPIPV	20%	25%	91%	95%	8%	19%
Env 47	VTVYYGVPVWK	59%	87%	95%	100%	61%	92%
Pol 929	QMAVFIHNFK	84%	98%	100%	100%	94%	97%
Pol 98	VTIKIGGQLK	11%	71%	59%	91%	2%	89%
Pol 971	KIQNFRVYYR	80%	86%	91%	95%	79%	89%
Pol 347	AIFQSSMTK	53%	75%	77%	82%	44%	79%
Pol 722	KVYLAWVPAHK	14%	97%	82%	95%	3%	97%
Env 61	TTLFCASDAK	72%	89%	90%	100%	69%	92%
Nef 94	FPVRPQVPL	81%	93%	77%	95%	82%	94%
Gag 545	YPLASLRSLF	7%	29%	45%	95%	0%	0%
Rev 75	VPLQLPPL	44%	78%	68%	77%	27%	79%
Env 259	IPIHYCAPA	74%	95%	45%	95%	79%	97%
Gag 237	HPVHAGPIA	27%	54%	68%	95%	44%	94%
Pol 893	IPYNPQSQGVV	92%	96%	82%	95%	240%	97%
Env 250	CPKVSFEPI	45%	91%	77%	100%	45%	97%
	Mean	**54%**	**81%**	**77%**	**93%**	**59%**	**84%**
	n=	**167**		**22**		**62**	

Table 8. HIV Peptides predicted to be immunologically conserved

Protein	Sequence	Conserved Epitopes*	Sequence Distribution					
				Subtype				
			All	A	B	C	D	G
Pol 498	ILKEPVHGV	**ILKEPVHGV**	104	1	17	46	2	2
		ILREPVHGV	12			5		1
		ILKEPVHGA	10			2	1	
		ILKDPVHGV	8	5				
		KLKEPVHGV	3					
		ILKDPVHGA	2	2				
		ILKNPVHGV	2					
Gag 386	VLAEAMSQV	**VLAEAMSQA**	67	2	1	36	3	3
		VLAEAMSQV	54	7	15	3	1	
		VLAEAMSQT	11			9		
		VLAEAMSHA	6			4		
		ILAEAMSQV	5		3			
		ILAEAMSQA	3			2		
		VLAEAMSHV	2					
Pol 448	KLVGKLNWA	**KLVGKLNWA**	158	9	21	59	3	3
		KLIGKLNWA	1					
Env 134	KLTPLCVTL	**KLTPLCVTL**	134	8	19	55		
		QLTPLCVTL	5	2	1			
		KLTPLCVAL	3					
		RLTPLCVTL	3			3		
		KITPLCVTL	2					
Vpr 62	RILQQLLFI	**RILQQLLFI**	86	1	15	28	4	3
		RILQQLLFV	21	2		2		
		RTLQQLLFI	10		2	4		
		RTLQQLLFV	10			1		
		RILQQLLFT	6			2		
		RMLQQLLFI	4		1	3		
		RVLQQLLFI	3			3		
Nef 221	LTFGWCFKL	**LTFGWCFKL**	82	8	17	29		3
		LTFGWCYKL	31	1	2	17		
		LTLGWCFKL	4			1		
Gag 271	MTNNPPIPV	**MTSNPPIPV**	60	3		24	4	1

		MTNNPPIPV	33		20	5		
		MTSNPPVPV	26	1		15		1
		MTGNPPIPV	15	5		1		
		MTGNPPVPV	9			5		
		MTNNPPVPV	6			6		
		MTANPPVPV	3			2		
Env 47	VTVYYGVPVWK	**VTVYYGVPVWK**	99	6	21	30	3	
		VTVYYGVPVWR	40	1		18		
		VTIYYGVPVWK	2					
Pol 929	QMAVFIHNFK	**QMAVFIHNFK**	153	10	22	58	4	3
		QMAVFVHNFK	3			1		
		QMAVFVHNYK	2					
Pol 98	VTIKIGGQLK	**VSIKVGGQIK**	30			30		
		VTIKIGGQLK	18		13	1		
		VTVKIGGQLK	11	1	1		1	
		VTVRIGGQLK	6	3				
		VSIKVGGQIR	6			6		
		VSIRVGGQIK	4			4		
		VTIRIGGQLK	3		2			
		VTVKIGGQLR	3	1				
		VTVKVGGQLK	3					
Pol 971	KIQNFRVYYR	**KIQNFRVYYR**	133	6	20	49	4	3
Pol 347	AIFQSSMTK	**AIFQSSMTK**	88	5	17	27	3	2
		AIFQCSMTK	19		2	5		
		AIFQSSMTR	13		1	11		1
		AIFQASMTK	9	1		1		
		SIFQSSMTK	9	3		6		
		AIFQYSMTK	4					
		AIFQSTMTK	2			1		
Pol 722	KVYLAWVPAHK	**KVYLSWVPAHK**	56	8		12	1	3
		RVYLSWVPAHK	55			41		
		KVYLAWVPAHK	23	1	18		3	
		KVYLTWVPAHK	5		2			
		KIYLSWVPAHK	5			3		
		RIYLSWVPAHK	5			4		
		KIYLAWVPAHK	2		1			
		QVYLTWVPAHK	2					
Env 61	TTLFCASDAK	**TTLFCASDAK**	121	9	19	41	4	1
		ATLFCASDAK	7			7		

		TILFCASDAK	6					
		PTLFCASDAK	2			1		
		TTLFCASDAR	2		2			
		TTLFCASEAK	2		1			
		ATLFCASDAR	2			2		
Nef 94	FPVRPQVPL	**FPVRPQVPL**	135	8	17	51	4	3
		FPVKPQVPL	9	1	3	2		
		FPVRPQVPV	4			2		
Gag 545	YPLASLRSLF	**EPLTSLKSLF***	22			21		
		YPLASLKSLF*	13		5		2	

*These two would not be predicted to XR. Would choose both to get maximal population coverage.

		YPLASLRSLF	11		10			
		YPLTSLKSLF	10		1		2	1
		YPPLASLKSL	10					
		YPLTSLRSLF	6		4			
		YPPLTSLKSL	6					
Rev 75	VPLQLPPL	**VPLQLPPL**	64	5	15	7	4	2
		VPLQLPPI	34	2	1	19		
		VPFQLPPI	26			23		
		VPFQLPPL	3		1			1
Env 259	IPIHYCAPA	**IPIHYCAPA**	124	8	10	49	3	2
		IPIHYCTPA	25	1	8	8		
		IPIHFCAPA	3		1	1	1	
Gag 237	HPVHAGPIA	**HPVHAGPIA**	39		15	21	1	
		HPVHAGPVA	34	1	3	27	2	
		HPVQAGPVA	12		3	6	1	
		HPIHAGPIA	2			2		
Pol 893	IPYNPQSQGVV	**IPYNPQSQGVV**	153	9	18	60	4	3
		IPYNPQSQGVI	5		3			
		IPYNPQSQGAV	2			1		
Env 250	CPKVSFEPI	**CPKVSFEPI**	50	5	17	3	2	0
		CPKVSFDPI	42			33		
		CPKVTFDPI	16			13		1
		CPKVTFEPI	13	3		1		1
		CPKISFDPI	9			5		
		CPKISFEPI	7		4		1	
		CPKVSWDPI	6					
		CPKVSFQPI	4		1			

* The preferred epitopes are shown in bold

Table 9. Predicted immunological conservation for a panel of HLA-A2 rest: peptides.

Source	Parent Sequence	HPV Strain	Variant Sequences	SEQ ID NO	Predicted Immunogenicity	Measured Immunogenicity (SU)*
HPV16.E7.86	TLGIVCPI	16	TLGIVCPI		+	103.7
		18	TLSFVCPW		-	
		31	SFGIVCPN		-	
		33	TVNIVCPT		-	
		45	TLSFVCPW		-	
		52	TLQVVCPG		-	
		56	ALTVTCPL		-	
		58	TCTIVCPS		-	
HPV31.E6.11	KLHELSSAL	16	KLPQLCTEL		-	
		18	KLPDLCTEL		-	
		31	KLHELSSAL		+	26.3
		33	TLHDLCQAL		-	
		45	KLPDLCTEL		-	
		52	TLHELCEVL		-	
		56	SLHHLSEVL		-	
		58	TLHDLCQAL		-	
HPV18/45.E6.13	KLPDLCTEL	16	KLPQLCTEL		+	15.7
		18	KLPDLCTEL		+	212.7
		31	KLHELSSAL		-	
		33	TLHDLCQAL		-	
		45	KLPDLCTEL		+	205.1
		52	TLHELCEVL		-	
		56	SLHHLSEVL		-	
		58	TLHDLCQAL		-	
HPV52.E6.18	VLEESVHEI	16	ELQTTIHDI		-	
		18	ELNTSLQDI		-	
		31	ALEIPYDEL		-	
		33	ALETTIHNI		-	
		45	ELNTSLQDV		-	
		52	VLEESVHEI		+	64.1
		56	VLEIPLIDL		-	
		58	ALETSVHEI		-	
HPV18.E6.47	FAFKDLFVV	16	FAFRDLCIV		-	
		18	FAFKDLFVV		+	350.6
		31	FAFTDLTIV		-	
		33	FAFADLTVV		-	31.4
		45	FAFKDLCIV		-	176.9
		52	FLFTDLRIV		-	
		56	FACTELKLV		-	
		58	FVFADLRIV		-	7.7
HPV31.E6.45	FAFTDLTIV	16	FAFRDLCIV		-	

		18	FAFKDLFVV		-	
		31	FAFTDLTIV		+	20.7
		33	FAFADLTVV		+	11.6
		45	FAFKDLCIV		-	
		52	FLFTDLRIV		-	
		56	FACTELKLV		-	
		58	FVFADLRIV		-	
HPV52.E6.45	FLFTDLRIV	16	FAFRDLCIV		-	
		18	FAFKDLFVV		-	
		31	FAFTDLTIV		-	
		33	FAFADLTVV		-	
		45	FAFKDLCIV		-	
		52	FLFTDLRIV		+	421.4
		56	FACTELKLV		-	57.5
		58	FVFADLRIV		+	94.1
HPV58.E6.45	FVFADLRIV	16	FAFRDLCIV		-	
		18	FAFKDLFVV		-	
		31	FAFTDLTIV		-	
		33	FAFADLTVV		-	
		45	FAFKDLCIV		-	
		52	FLFTDLRIV		+	13.3
		56	FACTELKLV		-	21.0
		58	FVFADLRIV		+	62.8
HPV18.E7.7	TLQDIVLHL	16	TLHEYMLDL		-	
		18	TLQDIVLHL		+	99.0
		31	TLQDYVLDL		-	
		33	TLKEYVLDL		-	
		45	TLQEIVLHL		+	
		52	TIKDYILDL		-	
		56	TLQDVVLEL		+	38.0
		58	TLREYILDL		-	
HPV16.E7.82	LLMGTLGIV	16	LLMGTLGIV		+	518.5
		18	LFLNTLSFV		-	
		31	LLMGSFGIV		+	90.1
		33	LLMGTVNIV		-	
		45	LFLSTLSFV		+	
		52	MLLGTLQVV		-	
		56	LLMGALTVT		+	
		58	LLMGTCTIV		-	
HPV33.E7.81	LLMGTVNIV	16	LLMGTLGIV		-	
		18	LFLNTLSFV		-	
		31	LLMGSFGIV		-	
		33	LLMGTVNIV		+	179.4
		45	LFLSTLSFV		-	
		52	MLLGTLQVV		+	
		56	LLMGALTVT		-	20.8
		58	LLMGTCTIV		-	
HPV52.E7.84	MLLGTLQVV	16	LLMGTLGIV		-	

		18	LFLNTLSFV		-	
		31	LLMGSFGIV		-	
		33	LLMGTVNIV		+	
		45	LFLSTLSFV		-	
		52	MLLGTLQVV		+	99.8
		56	LLMGALTVT		-	
		58	LLMGTCTIV		-	
HPV56.E7.89	LLMGALTVT	16	LLMGTLGIV		-	
		18	LFLNTLSFV		-	
		31	LLMGSFGIV		-	
		33	LLMGTVNIV		+	
		45	LFLSTLSFV		-	
		52	MLLGTLQVV		-	
		56	LLMGALTVT		+	263.5
		58	LLMGTCTIV		-	43.6

* Immunogenicity was measured for all variants. Only the positive responses are shown in the table. All other responses were negative.

Table 10. 167 HIV-1 Variants

SEQ ID NO	Sequence Designation	Name	Accession Number	SubType	Country
	A.UG.92UG037_U51190	92UG037	U51190	A	UG
	A.BY.97BL006_AF1932	97BL006	AF193275	A	BY
	A.KE.Q23_AF004885	Q23	AF004885	A	KE
	A.SE.SE6594_AF06967	SE6594	AF069672	A	SE
	A.SE.SE7253_AF06967	SE7253	AF069670	A	SE
	A.SE.SE7535_AF06967	SE7535	AF069671	A	SE
	A.SE.SE8538_AF06966	SE8538	AF069669	A	SE
	A.SE.SE8891_AF06967	SE8891	AF069673	A	SE
	A.UG.U455_M62320	U455	M62320	A	UG
	A.SE.UGSE8131_AF107	UGSE8131	AF107771	A	SE
	A2.CY.94CY017.41_AF	94CY017.41	AF286237	A2	CY
	A2.CD.97CDKTB48_AF2	97CDKTB48	AF286238	A2	CD
	A2D.-.97KR004_AF286	97KR004	AF286239	A2D	KR
	A2G.CD.97CDKP58_AF3	97CDKP58	AF316544	A2G	CD
	AC.IN.21301_AF06715	21301	AF067156	AC	IN
	AC.RW.92RW009_U8882	92RW009	U88823	AC	RW
	AC.SE.SE9488_AF0714	SE9488	AF071474	AC	SE
	ACD.SE.SE8603_AF075	SE8603	AF075702	ACD	SE
	ACG.BE.VI1035_AJ276	VI1035	AJ276595	ACG	BE
	AD.SE.SE6954_AF0757	SE6954	AF075701	AD	SE
	AD.SE.SE7108_AF0714	SE7108	AF071473	AD	SE
	ADHK.NO.97NOGIL3_AJ	97NOGIL3	AJ237565	ADHK	NO
	ADK.CD.MAL_X04415	MAL	X04415	ADK	CD
	AG.NG.92NG003_U8882	92NG003	U88825	AG	NG
	AG.BE.VI1197_AJ2765	VI1197	AJ276596	AG	BE
	AGHU.GA.VI354_AF076	VI354	AF076474	AGHU	GA
	AGU.CD.Z321_U76035	Z321	U76035	AGU	CD
	AJ.BW.BW2117_AF1921	BW2117	AF192135	AJ	BW
	B.NL.3202A21_U34604	3202A21	U34604	B	NL
	B.US.BC_L02317	BC	L02317	B	US
	B.GB.CAM1_D10112	CAM1	D10112	B	GB
	B.DE.D31_U43096	D31	U43096	B	DE
	B.US.DH123_AF069140	DH123	AF069140	B	US
	B.GB.GB8.C1_Y13716	GB8	AJ271445	B	GB
	B.DE.HAN_U43141	HAN	U43141	B	DE
	B.FR.HXB2_K03455	HXB2	K03455	B	FR
	B.US.JRCSF_M38429	JRCSF	M38429	B	US
	B.GB.MANC_U23487	MANC	U23487	B	GB
	B.US.MNCG_M17449	MNCG	M17449	B	US
	B.GA.OYI,_M26727	OYI	M26727	B	GA
	B.US.P896_U39362	P896	M96155	B	US
	B.US.RF_M17451	RF	M17451	B	US
	B.CN.RL42_U71182	RL42	U71182	B	CN
	B.US.SF2_K02007	SF2	K02007	B	US
	B.TW.TWCYS_AF086817	TWCYS	AF086817	B	TW
	B.AU.VH_AF146728	VH	AF146728	B	AU
	B.US.WEAU160_U21135	WEAU160	U21135	B	US
	B.KR.WK_AF224507	WK	AF224507	B	KR
	B.US.WR27_U26546	WR27	U26546	B	US

B.US.YU2_M93258	YU2	M93258	B	US
BF1.BR.93BR029.4_AF	93BR029.4	AF005495	BF1	BR
C.BR.92BR025_U52953	92BR025	U52953	C	BR
C.IN.93IN101_AB0238	93IN101	AB023804	C	IN
C.IN.93IN904_AF0671	93IN904	AF067157	C	IN
C.IN.93IN999_AF0671	93IN999	AF067154	C	IN
C.IN.94IN11246_AF06	94IN11246	AF067159	C	IN
C.IN.95IN21068_AF06	95IN21068	AF067155	C	IN
C.BW.96BW0402_AF110	96BW0402	AF110962	C	BW
C.BW.96BW1210_AF110	96BW1210	AF110972	C	BW
C.BW.96BW15B03_AF11	96BW15B03	AF110973	C	BW
C.ET.ETH2220_U46016	ETH2220	U46016	C	ET
C.BW.96BW11B01_AF11	96BW11	AF110969	C	BW
C.BW.00BW0762.1_AF44	00BW0762.1	AF443088	C	BW
C.BW.00BW0768.20_AF44	00BW0768.20	AF443089	C	BW
C.BW.00BW0874.21_AF44	00BW0874.21	AF443090	C	BW
C.BW.00BW1471.27_AF44	00BW1471.27	AF443091	C	BW
C.BW.00BW1616.2_AF44	00BW1616.2	AF443092	C	BW
C.BW.00BW1686.8_AF44	00BW1686.8	AF443093	C	BW
C.BW.00BW1759.3_AF44	00BW1759.3	AF443094	C	BW
C.BW.00BW1773.2_AF44	00BW1773.2	AF443095	C	BW
C.BW.00BW1783.5_AF44	00BW1783.5	AF443096	C	BW
C.BW.00BW1795.6_AF44	00BW1795.6	AF443097	C	BW
C.BW.00BW1811.3_AF44	00BW1811.3	AF443098	C	BW
C.BW.00BW1859.5_AF44	00BW1859.5	AF443099	C	BW
C.BW.00BW1880.2_AF44	00BW1880.2	AF443100	C	BW
C.BW.00BW1921.13_AF44	00BW1921.13	AF443101	C	BW
C.BW.00BW2036.1_AF44	00BW2036.1	AF443102	C	BW
C.BW.00BW2063.6_AF44	00BW2063.6	AF443103	C	BW
C.BW.00BW2087.2_AF44	00BW2087.2	AF443104	C	BW
C.BW.00BW2127.214_AF44	00BW2127.214	AF443105	C	BW
C.BW.00BW2128.3_AF44	00BW2128.3	AF443106	C	BW
C.BW.00BW2276.7_AF44	00BW2276.7	AF443107	C	BW
C.BW.00BW3819.3_AF44	00BW3819.3	AF443108	C	BW
C.BW.00BW3842.8_AF44	00BW3842.8	AF443109	C	BW
C.BW.00BW3871.3_AF44	00BW3871.3	AF443110	C	BW
C.BW.00BW3876.9_AF44	00BW3876.9	AF443111	C	BW
C.BW.00BW3886.8_AF44	00BW3886.8	AF443112	C	BW
C.BW.00BW3891.6_AF44	00BW3891.6	AF443113	C	BW
C.BW.00BW3970.2_AF44	00BW3970.2	AF443114	C	BW
C.BW.00BW5031.1_AF44	00BW5031.1	AF443115	C	BW
C.BW.96BW01B21_AF11	96BW01B21	AF110960	C	BW
C.BW.96BW0407_AF11	96BW0407	AF110963	C	BW
C.BW.96BW0502_AF11	96BW0502	AF110967	C	BW
C.BW.96BW06.J4_AF29	96BW06.J4	AF290028	C	BW
C.BW.96BW11.06_AF11	96BW11.06	AF110970	C	BW
C.BW.96BW1210_AF11	96BW1210	AF110972	C	BW
C.BW.96BW15B03_AF11	96BW15B03	AF110973	C	BW
C.BW.96BW16.26_AF11	96BW16.26	AF110978	C	BW
C.BW.96BW17A09_AF11	96BW17A09	AF110979	C	BW
C.BW.96BWMO1.5_AF44	96BWMO1.5	AF443074	C	BW
C.BW.96BWMO3.2_AF44	96BWMO3.2	AF443075	C	BW
C.BW.98BWMC12.2_AF44	98BWMC12.2	AF443076	C	BW
C.BW.98BWMC13.4_AF44	98BWMC13.4	AF443077	C	BW
C.BW.98BWMC14.a3_AF44	98BWMC14.a3	AF443078	C	BW

C.BW.98BWMO14.10_AF44	98BWMO14.10	AF443079	C	BW
C.BW.98BWMO18.d5_AF44	98BWMO18.d5	AF443080	C	BW
C.BW.98BWMO36.a5_AF44	98BWMO36.a5	AF443081	C	BW
C.BW.98BWMO37.d5_AF44	98BWMO37.d5	AF443082	C	BW
C.BW.99BW3932.12_AF44	99BW3932.12	AF443083	C	BW
C.BW.99BW4642.4_AF44	99BW4642.4	AF443084	C	BW
C.BW.99BW4745.8_AF44	99BW4745.8	AF443085	C	BW
C.BW.99BW4754.7_AF44	99BW4754.7	AF443086	C	BW
C.BW.99BWMC16.8_AF44	99BWMC16.8	AF443087	C	BW
CRF01_AE.CF.90CF11697_	90CF11697	AF197340	CRF01_AE	CF
CRF01_AE.CF.90CF402_U5	90CF402	U51188	CRF01_AE	CF
CRF01_AE.CF.90CF4071_A	90CF4071	AF197341	CRF01_AE	CF
CRF01_AE.TH.93TH057_AF	93TH057	AF197338	CRF01_AE	TH
CRF01_AE.TH.93TH065_AF	93TH065	AF197339	CRF01_AE	TH
CRF01_AE.TH.93TH253_U5	93TH253	U51189	CRF01_AE	TH
CRF01_AE.TH.95TNIH047_	95TNIH047	AB032741	CRF01_AE	TH
CRF01_AE.TH.CM240_U547	CM240	U54771	CRF01_AE	TH
CRF01_AE.TH.TH022_AB03	TH022	AB032740	CRF01_AE	TH
CRF02_AG.SN.98SEMP1211	98SEMP1211	AJ251056	CRF02_AG	SN
CRF02_AG.FR.DJ263_AF06	DJ263	AF063223	CRF02_AG	FR
CRF02_AG.FR.DJ264_AF06	DJ264	AF063224	CRF02_AG	FR
CRF02_AG.GH.G829_AF184	G829	AF184155	CRF02_AG	GH
CRF02_AG.NG.IBNG_L3910	IBNG	L39106	CRF02_AG	NG
CRF02_AG.SE.SE7812_AF1	SE7812	AF107770	CRF02_AG	SE
CRF03_AB.RU.KAL153-2_A	KAL153-2	AF193276	CRF03_AB	RU
CRF03_AB.RU.RU98001_AF	RU98001	AF193277	CRF03_AB	RU
CRF04_cpx.CY.94CY032-3	94CY032-3	AF049337	CRF04_cpx	CY
CRF04_cpx.GR.97PVCH_AF	97PVCH	AF119820	CRF04_cpx	GR
CRF04_cpx.GR.97PVMY_AF	97PVMY	AF119819	CRF04_cpx	GR
CRF05_DF.BE.VI1310_AF1	VI1310	AF193253	CRF05_DF	BE
CRF05_DF.BE.VI961_AF07	VI961	AF076998	CRF05_DF	BE
CRF06_cpx.ML.95ML127_A	95ML127	AJ288982	CRF06_cpx	ML
CRF06_cpx.ML.95ML84_AJ	95ML84	AJ245481	CRF06_cpx	ML
CRF06_cpx.SN.97SE1078_	97SE1078	AJ288981	CRF06_cpx	SN
CRF06_cpx.AU.BFP90_AF0	BFP90	AF064699	CRF06_cpx	AU
CRF11_cpx.CM.97CM-MP81	97CM-MP818	AJ291718	CRF11_cpx	CM
CRF11_cpx.GR.GR17_AF17	GR17	AF179368	CRF11_cpx	GR
D.CD.84ZR085_U88822	84ZR085	U88822	D	CD
D.UG.94UG1141_U8882	94UG1141	U88824	D	UG
D.CD.ELI_K03454	ELI	K03454	D	CD
D.CD.NDK_M27323	NDK	M27323	D	CD
F1.BR.93BR020.1_AF0	93BR020.1	AF005494	F1	BR
F1.FI.FIN9363_AF075	FIN9363	AF075703	F1	FI
F1.FR.MP411_AJ24923	MP411	AJ249238	F1	FR
F1.BE.VI850_AF07733	VI850	AF077336	F1	BE
F2.CM.MP257_AJ24923	MP257	AJ249237	F2	CM
F2KU.BE.VI1126_AF07	VI1126	AF076475	F2KU	BE
G.NG.92NG083_U88826	92NG083	U88826	G	NG
G.BE.DRCBL_AF084936	DRCBL	AF084936	G	BE
G.SE.SE6165_AF06164	SE6165	AF061642	G	SE
H.CF.90CF056_AF0054	90CF056	AF005496	H	CF
H.BE.VI991_AF190127	VI991	AF190127	H	BE
H.BE.VI997_AF190128	VI997	AF190128	H	BE
J.SE.SE7022_AF08239	SE7022	AF082395	J	SE
J.SE.SE7887_AF08239	SE7887	AF082394	J	SE

K.CD.EQTB11C_AJ2492	EQTB11C	AJ249235	K	CD
K.CM.MP535_AJ249239	MP535	AJ249239	K	CM
N.CM.YBF30_AJ006022	YBF30	AJ006022	N	CM
O.SN.99SE-MP1299_ZX	SEMP1299	AJ302646	O	SN
O.SN.99SE-MP1300_ZX	SEMP1300	AJ302647	O	SN
O.CM.ANT70_L20587	ANT70	L20587	O	CM
O.CM.MVP5180_L20571	MVP5180	L20571	O	CM
U.CD.-.83CD0031	83CD0031	AF286236	U	CD

Table 11. HIV Gag Sequence Alignment

GCG Multiple Sequence File.
Written by Omiga 1.1

Name:	00BW0762_1	Len:	556	Check:	2513	Weight:	1.00
Name:	00BW0768_2	Len:	556	Check:	8965	Weight:	1.00
Name:	00BW0874_2	Len:	556	Check:	9574	Weight:	1.00
Name:	00BW1471_2	Len:	556	Check:	5395	Weight:	1.00
Name:	00BW1616_2	Len:	556	Check:	4692	Weight:	1.00
Name:	00BW1686_8	Len:	556	Check:	7822	Weight:	1.00
Name:	00BW1759_3	Len:	556	Check:	7777	Weight:	1.00
Name:	00BW1773_2	Len:	556	Check:	9727	Weight:	1.00
Name:	00BW1783_5	Len:	556	Check:	9681	Weight:	1.00
Name:	00BW1795_6	Len:	556	Check:	9667	Weight:	1.00
Name:	00BW1811_3	Len:	556	Check:	4422	Weight:	1.00
Name:	00BW1859_5	Len:	556	Check:	7320	Weight:	1.00
Name:	00BW1880_2	Len:	556	Check:	1603	Weight:	1.00
Name:	00BW1921_1	Len:	556	Check:	883	Weight:	1.00
Name:	00BW2036_1	Len:	556	Check:	2591	Weight:	1.00
Name:	00BW2063_6	Len:	556	Check:	5152	Weight:	1.00
Name:	00BW2087_2	Len:	556	Check:	5183	Weight:	1.00
Name:	00BW2127_2	Len:	556	Check:	5469	Weight:	1.00
Name:	00BW2128_3	Len:	556	Check:	9621	Weight:	1.00
Name:	00BW2276_7	Len:	556	Check:	4153	Weight:	1.00
Name:	00BW3819_3	Len:	556	Check:	4227	Weight:	1.00
Name:	00BW3842_8	Len:	556	Check:	9312	Weight:	1.00
Name:	00BW3871_3	Len:	556	Check:	501	Weight:	1.00
Name:	00BW3876_9	Len:	556	Check:	773	Weight:	1.00
Name:	00BW3886_8	Len:	556	Check:	2351	Weight:	1.00
Name:	00BW3891_6	Len:	556	Check:	129	Weight:	1.00
Name:	00BW3970_2	Len:	556	Check:	8768	Weight:	1.00
Name:	00BW5031_1	Len:	556	Check:	3966	Weight:	1.00
Name:	96BW01B21	Len:	556	Check:	602	Weight:	1.00
Name:	96BW0407	Len:	556	Check:	9836	Weight:	1.00
Name:	96BW0502	Len:	556	Check:	6402	Weight:	1.00
Name:	96BW06_J4	Len:	556	Check:	254	Weight:	1.00
Name:	96BW11_06	Len:	556	Check:	6801	Weight:	1.00
Name:	96BW1210	Len:	556	Check:	6016	Weight:	1.00
Name:	96BW15B03	Len:	556	Check:	6072	Weight:	1.00
Name:	96BW16_26	Len:	556	Check:	9409	Weight:	1.00
Name:	96BW17A09	Len:	556	Check:	2723	Weight:	1.00
Name:	96BWMO1_5	Len:	556	Check:	5051	Weight:	1.00
Name:	96BWMO3_2	Len:	556	Check:	496	Weight:	1.00
Name:	98BWMC12_2	Len:	556	Check:	1164	Weight:	1.00
Name:	98BWMC13_4	Len:	556	Check:	4961	Weight:	1.00
Name:	98BWMC14_a	Len:	556	Check:	7351	Weight:	1.00
Name:	98BWMO14_1	Len:	556	Check:	288	Weight:	1.00
Name:	98BWMO18_d	Len:	556	Check:	6836	Weight:	1.00
Name:	98BWMO36_a	Len:	556	Check:	4386	Weight:	1.00
Name:	98BWMO37_d	Len:	556	Check:	6900	Weight:	1.00
Name:	99BW3932_1	Len:	556	Check:	292	Weight:	1.00
Name:	99BW4642_4	Len:	556	Check:	1347	Weight:	1.00
Name:	99BW4745_8	Len:	556	Check:	7980	Weight:	1.00
Name:	99BW4754_7	Len:	556	Check:	9892	Weight:	1.00
Name:	99BWMC16_8	Len:	556	Check:	4279	Weight:	1.00
Name:	A2_CD_97CD	Len:	556	Check:	9920	Weight:	1.00
Name:	A2_CY_94CY	Len:	556	Check:	3070	Weight:	1.00
Name:	A2D___97KR	Len:	556	Check:	6350	Weight:	1.00
Name:	A2G_CD_97C	Len:	556	Check:	856	Weight:	1.00
Name:	A_BY_97BL0	Len:	556	Check:	9315	Weight:	1.00
Name:	A_KE_Q23_A	Len:	556	Check:	2442	Weight:	1.00

Name: A_SE_SE659 Len: 556 Check: 8612 Weight: 1.00
Name: A_SE_SE725 Len: 556 Check: 8315 Weight: 1.00
Name: A_SE_SE753 Len: 556 Check: 2915 Weight: 1.00
Name: A_SE_SE853 Len: 556 Check: 9112 Weight: 1.00
Name: A_SE_SE889 Len: 556 Check: 8732 Weight: 1.00
Name: A_SE_UGSE8 Len: 556 Check: 8696 Weight: 1.00
Name: A_UG_92UG0 Len: 556 Check: 6290 Weight: 1.00
Name: A_UG_U455_ Len: 556 Check: 164 Weight: 1.00
Name: AC_IN_2130 Len: 556 Check: 8482 Weight: 1.00
Name: AC_RW_92RW Len: 556 Check: 977 Weight: 1.00
Name: AC_SE_SE94 Len: 556 Check: 8752 Weight: 1.00
Name: ACD_SE_SE8 Len: 556 Check: 9655 Weight: 1.00
Name: ACG_BE_VI1 Len: 556 Check: 3777 Weight: 1.00
Name: AD_SE_SE69 Len: 556 Check: 732 Weight: 1.00
Name: AD_SE_SE71 Len: 556 Check: 8506 Weight: 1.00
Name: ADHK_NO_97 Len: 556 Check: 5257 Weight: 1.00
Name: ADK_CD_MAL Len: 556 Check: 5301 Weight: 1.00
Name: AG_BE_VI11 Len: 556 Check: 1610 Weight: 1.00
Name: AG_NG_92NG Len: 556 Check: 4188 Weight: 1.00
Name: AGHU_GA_VI Len: 556 Check: 8242 Weight: 1.00
Name: AGU_CD_Z32 Len: 556 Check: 2601 Weight: 1.00
Name: AJ_BW_BW21 Len: 556 Check: 8389 Weight: 1.00
Name: B_AU_VH_AF Len: 556 Check: 9166 Weight: 1.00
Name: B_CN_RL42_ Len: 556 Check: 3865 Weight: 1.00
Name: B_DE_D31_U Len: 556 Check: 9464 Weight: 1.00
Name: B_DE_HAN_U Len: 556 Check: 6999 Weight: 1.00
Name: B_FR_HXB2_ Len: 556 Check: 141 Weight: 1.00
Name: B_GA_OYI__ Len: 556 Check: 7578 Weight: 1.00
Name: B_GB_CAM1_ Len: 556 Check: 8023 Weight: 1.00
Name: B_GB_GB8_A Len: 556 Check: 2675 Weight: 1.00
Name: B_GB_MANC_ Len: 556 Check: 8961 Weight: 1.00
Name: B_KR_WK_AF Len: 556 Check: 7372 Weight: 1.00
Name: B_NL_3202A Len: 556 Check: 7117 Weight: 1.00
Name: B_TW_TWCYS Len: 556 Check: 8767 Weight: 1.00
Name: B_US_BC_L0 Len: 556 Check: 7091 Weight: 1.00
Name: B_US_DH123 Len: 556 Check: 5049 Weight: 1.00
Name: B_US_JRCSF Len: 556 Check: 975 Weight: 1.00
Name: B_US_MNCG_ Len: 556 Check: 688 Weight: 1.00
Name: B_US_P896_ Len: 556 Check: 8809 Weight: 1.00
Name: B_US_RF_M1 Len: 556 Check: 9306 Weight: 1.00
Name: B_US_SF2_K Len: 556 Check: 9799 Weight: 1.00
Name: B_US_WEAU1 Len: 556 Check: 9636 Weight: 1.00
Name: B_US_WR27_ Len: 556 Check: 3349 Weight: 1.00
Name: B_US_YU2_M Len: 556 Check: 8828 Weight: 1.00
Name: BF1_BR_93B Len: 556 Check: 7935 Weight: 1.00
Name: C_BR_92BR0 Len: 556 Check: 4429 Weight: 1.00
Name: C_BW_96BW0 Len: 556 Check: 1550 Weight: 1.00
Name: C_BW_96BW1 Len: 556 Check: 7158 Weight: 1.00
Name: C_BW_96BW1 Len: 556 Check: 6016 Weight: 1.00
Name: C_BW_96BW1 Len: 556 Check: 6072 Weight: 1.00
Name: C_ET_ETH22 Len: 556 Check: 4314 Weight: 1.00
Name: C_IN_93IN1 Len: 556 Check: 6959 Weight: 1.00
Name: C_IN_93IN9 Len: 556 Check: 9362 Weight: 1.00
Name: C_IN_93IN9 Len: 556 Check: 3298 Weight: 1.00
Name: C_IN_94IN1 Len: 556 Check: 6744 Weight: 1.00
Name: C_IN_95IN2 Len: 556 Check: 8559 Weight: 1.00
Name: CRF01_AE_C Len: 556 Check: 4763 Weight: 1.00
Name: CRF01_AE_C Len: 556 Check: 4315 Weight: 1.00
Name: CRF01_AE_C Len: 556 Check: 3920 Weight: 1.00
Name: CRF01_AE_T Len: 556 Check: 5074 Weight: 1.00
Name: CRF01_AE_T Len: 556 Check: 4026 Weight: 1.00

```
Name: CRF01_AE_T        Len:   556  Check: 2981  Weight:   1.00
Name: CRF01_AE_T        Len:   556  Check: 6648  Weight:   1.00
Name: CRF01_AE_T        Len:   556  Check: 6948  Weight:   1.00
Name: CRF01_AE_T        Len:   556  Check: 6129  Weight:   1.00
Name: CRF02_AG_F        Len:   556  Check: 8134  Weight:   1.00
Name: CRF02_AG_F        Len:   556  Check: 8925  Weight:   1.00
Name: CRF02_AG_G        Len:   556  Check: 9991  Weight:   1.00
Name: CRF02_AG_N        Len:   556  Check: 9388  Weight:   1.00
Name: CRF02_AG_S        Len:   556  Check: 9371  Weight:   1.00
Name: CRF02_AG_S        Len:   556  Check: 3289  Weight:   1.00
Name: CRF03_AB_R        Len:   556  Check: 5554  Weight:   1.00
Name: CRF03_AB_R        Len:   556  Check: 5307  Weight:   1.00
Name: CRF04_cpx_        Len:   556  Check: 5488  Weight:   1.00
Name: CRF04_cpx_        Len:   556  Check: 6046  Weight:   1.00
Name: CRF04_cpx_        Len:   556  Check: 3807  Weight:   1.00
Name: CRF05_DF_B        Len:   556  Check: 6702  Weight:   1.00
Name: CRF05_DF_B        Len:   556  Check: 5563  Weight:   1.00
Name: CRF06_cpx_        Len:   556  Check: 8164  Weight:   1.00
Name: CRF06_cpx_        Len:   556  Check: 1975  Weight:   1.00
Name: CRF06_cpx_        Len:   556  Check: 1486  Weight:   1.00
Name: CRF06_cpx_        Len:   556  Check: 4050  Weight:   1.00
Name: CRF11_cpx_        Len:   556  Check: 1515  Weight:   1.00
Name: CRF11_cpx_        Len:   556  Check: 9474  Weight:   1.00
Name: D_CD_84ZR0        Len:   556  Check: 4739  Weight:   1.00
Name: D_CD_ELI_K        Len:   556  Check: 8533  Weight:   1.00
Name: D_CD_NDK_M        Len:   556  Check: 6100  Weight:   1.00
Name: D_UG_94UG1        Len:   556  Check: 7681  Weight:   1.00
Name: F1_BE_VI85        Len:   556  Check: 8540  Weight:   1.00
Name: F1_BR_93BR        Len:   556  Check: 241  Weight:   1.00
Name: F1_FI_FIN9        Len:   556  Check: 7510  Weight:   1.00
Name: F1_FR_MP41        Len:   556  Check: 9846  Weight:   1.00
Name: F2_CM_MP25        Len:   556  Check: 1247  Weight:   1.00
Name: F2KU_BE_VI        Len:   556  Check: 2186  Weight:   1.00
Name: G_BE_DRCBL        Len:   556  Check: 8049  Weight:   1.00
Name: G_NG_92NG0        Len:   556  Check: 254  Weight:   1.00
Name: G_SE_SE616        Len:   556  Check: 6407  Weight:   1.00
Name: H_BE_VI991        Len:   556  Check: 808  Weight:   1.00
Name: H_BE_VI997        Len:   556  Check: 4749  Weight:   1.00
Name: H_CF_90CF0        Len:   556  Check: 4718  Weight:   1.00
Name: J_SE_SE702        Len:   556  Check: 2564  Weight:   1.00
Name: J_SE_SE788        Len:   556  Check: 2692  Weight:   1.00
Name: K_CD_EQTB1        Len:   556  Check: 6586  Weight:   1.00
Name: K_CM_MP535        Len:   556  Check: 3803  Weight:   1.00
Name: N_CM_YBF30        Len:   556  Check: 8603  Weight:   1.00
Name: O_CM_ANT70        Len:   556  Check: 365  Weight:   1.00
Name: O_CM_MVP51        Len:   556  Check: 9160  Weight:   1.00
Name: O_SN_99SE_        Len:   556  Check: 9380  Weight:   1.00
Name: O_SN_99SE_        Len:   556  Check: 8796  Weight:   1.00
Name: U_CD___83C        Len:   556  Check: 4633  Weight:   1.00

//

                1                                                    50
00BW0762_1   MGARASILRG .EKLDKWEKI RLRPGGRKHY MIKHIVWASR ELERFALNPG
00BW0768_2   MGARASVLRG .EKLDKWEKI RLRPGGKKHY MLKHLVWASR ELERFALNPG
00BW0874_2   MGARASILRG .GKLDTWEKI RLRPGGKKQY MIKHLVWASR ELERFALNPG
00BW1471_2   MGARASILRG .GKLDTWEKI RLRPGGKKHY MLKHLVWASR ELERFALNPG
00BW1616_2   MGARASILRG .GKLDEWEKI RLRPGGKKRY MMKHLVWASR ELERFALNPG
00BW1686_8   MGARASILRG .GKLDTWEKI RLRPGGKKHY MIKHLVWASR ELERFALNPG
00BW1759_3   MGARASILRG .GKLDKWERI RLRPGGKKHY MLKHLVWASR ELERFALNPG
00BW1773_2   MGASASILRG .GKLDKWEKI RLRPGGKKKY RLKHLVWASR ELERFALNSG
```

```
00BW1783_5   MGARASILRG .GKLDTWEKI RLRPGGKKHY MMKHLVWASR ELERFALNPG
00BW1795_6   MGARASILRG .GKLDKWEKI RLRPGGKKHY MMKHLVWASR ELERFALNPG
00BW1811_3   MGARASILRG .GKLDKWEKI RLRPGGKKHY MLKHLVWASR ELERFALNPG
00BW1859_5   MGARASVLKG .EKLDAWEKI RLRPGEKKGY MLKHLVWASR ELERFALNPG
00BW1880_2   MGARASILRG .GKLDKWERI RLRPGGKKQY MIKHLVWASR ELERFALNPG
00BW1921_1   MGARASILRG .GKLDTWEKI RLRPGGKKRY MLKHLVWASR ELERFALNPG
00BW2036_1   MGARASILRG .EKLDTWERI KLRPGGKKHY MLKHLIWTSR ELERFALNPD
00BW2063_6   MGARASILRG .GKLDKWEKI RLRPGGKKQY MIKHLVWASR ELERFALNPG
00BW2087_2   MGARASILRG .GKLDTWEKI KLRPGGKKSY KLKHLVWASR ELERFALNPG
00BW2127_2   MGARASILRG XEKLDEWEKI RLRPGGRKKY RLKHLVWASR ELERFALNPS
00BW2128_3   MGARASILRG .GQLDKWEKI RLRPGGKKHY MLKHLVWASR ELENFALNPG
00BW2276_7   MGARASVLKG .DKLDAWEKI KLRPGGKKHY MLKHLVWASG ELEKFALNPG
00BW3819_3   MGARASILRG .GKLDAWERT RLRPGGKKHY MLKHLVWASR ELERFALNPG
00BW3842_8   MGARASVLRG .EKLDTWERI KLRPGGKKHY RLKHLVWASR ELERFALNPG
00BW3871_3   MGARASILRG .GKLDTWEKI RLRPGGKKHY MLKHIVWASR ELERFALNPG
00BW3876_9   MGARASILKG .GKLDTWEKI RLRPGGKKHY MLKHLVWASR ELERFALNPG
00BW3886_8   MGARASILRG .GKLDKWEKI RLRPGGKKHY MVKHLVWASR ELERFALNPG
00BW3891_6   MGARASILRG .GKLDKWEKI RLRPGGKKCY MIKHIIWASR ELERFALNPG
00BW3970_2   MGARASILRG .GKLDAWERI KLRPGGKKKY MLKHLVWASR ELERFALNSG
00BW5031_1   MGARASILRG .GKLDRWEKI RLRPGGKKHY MLKHLVWASR ELERFALNPS
 96BW01B21   MGARASILRG .GKLDQWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPS
  96BW0407   MGARASILRG .GKLDAWERI RLRPGGKKCY MLKHLVWASR ELERFALNPG
  96BW0502   MGARASILRG .EKLDKWEKI RLRPGGKKCY MMKHLVWASR ELEKFALNPG
 96BW06_J4   MGARASILRG .GNLDTWEKI RLRPGGKKHY MLKHLVWASR ELEGFALNPG
 96BW11_06   MGARASILRG .GKLDKWEKI RLRPGGKKHY MLKHLVWASR ELERFALNPG
  96BW1210   MGARASILRG .EKLDTWEKI RLRPGGKKRY MIKHLVWASR ELERFALNPG
 96BW15B03   MGARASISRG .EKLDTWEKI RSRPGGKKCY MLKHIVWASR ELENFALNPA
 96BW16_26   MGARASILRG .GKLDKWEKI RLRPGGKKRY MLKHLVWASR ELERFALNPG
 96BW17A09   MGARASILRG .GKLDTWEKI RLRPGGKKHY MLKHLVWASR ELERFALNPG
 96BWMO1_5   MGARASILRR .GKLDKWEKI RLRPGGKKRY MIKHLVWASR ELERFALNPG
 96BWMO3_2   MGARASILRG .EKLDKWEKI RLRPGGKKHY MLKHIVWASR ELERYALNPG
98BWMC12_2   MGARASILRG .EKLDTWEKI RLRPGGKKQY RIKHLVWASR ELEKFALNPG
98BWMC13_4   MGARASILRG .GKLDKWEKI RLRPGGKKHY MMKHLVWASR ELDRFALNSG
98BWMC14_a   MGARASILRG .GKLDKWEKI RLRPGGKKHY MMKHLVWASR ELGRFALNPG
98BWMO14_1   MGARASILRG .GKLDKWERI RLRPGGKKQY MLKHLVWASR ELERFALNPG
98BWMO18_d   MGARASILRG .GKLDTWERI KLRPGGKKHY MMKHLVWASR ELERFALNPG
98BWMO36_a   MGARASILRG .GKLDTWEKI RLRPGGKKPY MLKHLVWASR ELERFALNPS
98BWMO37_d   MGARASILRG .GKLDAWEKI RLRPGGKKHY MLKHLVWASR ELERFALNPG
99BW3932_1   MGARASILRG .GKLDEWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPE
99BW4642_4   MGARASVLKG .EKLDTWEKI RLRPGGRKHY MLKHLVWASR ELERFALNPG
99BW4745_8   MGARASILRG .GKLGNWERI KLRPGGRKTY MLKHLVWASR ELERFALNPS
99BW4754_7   MGARASILRG .EKLDRWEKI WLRPGGKNHY MLKHLVWASR ELERFALNPA
99BWMC16_8   MGARASILRG .GKLDTWEKI RLRPGGKKHY MIKHLVWASR ELERFALNPG
A2_CD_97CD   MGARASVLSG .GKLEAWEKI RLRPGGKKKY RLKHLVWASR ELEKFSINPS
A2_CY_94CY   MGARASILSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELEKFSINPG
A2D___97KR   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFAINPG
A2G_CD_97C   ..ARASILSG .GKLEAWEKI RLRPGGKKKY RLKHLVWASR ELEKFSINPG
A_BY_97BL0   XGARASVLSG .GKLDA.EKI RLRPXGKKKY RIKHLVWASR ELERFALNPG
A_KE_Q23_A   MGARASVLSG .GKFDAWEKI RLRPGGKKKY RMKHLIWASR ELDRFALNPS
A_SE_SE659   ...RASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPS
A_SE_SE725   ...RASVLSG .GKLDAWEKI RLRPGGKKKY RMKHLVWASR ELERFALNPS
A_SE_SE753   MGARASVLSG .GKLDAWEKI RLRPGGKKQY RLKHLVWASR ELERFALNPS
A_SE_SE853   ...RASVLSG .GRLDAWEKI RLRPGGKKKY RMKHLVWASR ELDRFALNPS
A_SE_SE889   ...RASKLSG .EKKDAWEKM RLRPGGKKKY KLKHMVWARR ELEKSALNPS
A_SE_UGSE8   MGARASVLSG .GKLDAWEKI RLRPGGNKKY RLKHLVWASR ELEKFALNPG
A_UG_92UG0   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPS
A_UG_U455_   MGARASVLSG .KKLDSWEKI RLRPGGNKKY RLKHLVWASR ELEKFTLNPG
AC_IN_2130   MGARASILRG .GKLDKWEKI RLRPGGKKHY MIKHLVWASR ELERFALNPG
AC_RW_92RW   MGARASILRG .GKLDAWEKI KLKPGGKKTY MMKHLVWASR ELERFALNPD
AC_SE_SE94   ...RASVLSG .GKLDAWEKI RLRPGGKKKY RMKHLVWASR ELDRFALNPS
ACD_SE_SE8   MGARASILSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELDRFALNPS
```

```
ACG_BE_VI1   MGARASVLTG .GKLDTWERI RLRPAGKKKY RMKLLVWASR ELERFAINPG
AD_SE_SE69   MGARASVLSG .ERLDEWEKI QLRPGGKKRY RLKHIVWASR ELERFALNPG
AD_SE_SE71   ...RASVLSG .GKLDAWEKI RLRPGGRKKY KLKHIVWASR ELERFALNPS
ADHK_NO_97   MGARASILSG .GKLDKWEKI RLRPGGKKQY RLKHLVWASR ELDRFALNPS
ADK_CD_MAL   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPG
AG_BE_VI11   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RMKHLVWASR ELERFALNPG
AG_NG_92NG   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RMKHLVWASR ELERFAINPG
AGHU_GA_VI   MGARASILSG .GKLDAWEKI RLRPGGKKKY QIKHIVWASR ELERFALNPD
AGU_CD_Z32   MGARASVLSG .GKLDKWEKI RLRPGGKKQY KLKHIVWASR ELERFALNPG
AJ_BW_BW21   MGARASVLSG .GKLDAWEKI RLRPGGKKQY RMKHLVWASR ELERFALNPG
B_AU_VH_AF   MGARASVLSG .GELDRWEKI RLRPGGKKTY KLKHIVWASR ELERFALNPG
B_CN_RL42_   MGARASVLSG .GQLDRWEKI RLRPGGKKKY RLKHLVWASR ELERFAVNPG
B_DE_D31_U   MGARASVLSG .GELDRWEKI RLRPGGKKKY RLKHLVWASR ELERFAVNPG
B_DE_HAN_U   MGARASVLSG .GELDKWEKI RLRPGGKKKY RLKHIVWASR ELERFAVNPG
B_FR_HXB2_   MGARASVLSG .GELDRWEKI RLRPGGKKKY QLKHIVWASR ELERFAVNPG
B_GA_OYI__   MGARASVLSG .GELDKWEKI RLRPGGKKKY KLKHIVWASR ELERFAVNPG
B_GB_CAM1_   MGARASVLSG .GELDKWEKI RLRPGGKKKY QLKHIVWASR ELERFAINPG
B_GB_GB8_A   MGARASVLSG .GELDRWEKI RLRPGGKKKY KLKHIVWASR ELERFAVNPG
B_GB_MANC_   MGARASVLSG .GKLDRWEKI RLRPGGKKKY RLKHVVWASR ELERFAVNPG
B_KR_WK_AF   MGARASILSG .GELDQWEKI RLRPGGKKKY KLKHIVWASR ELERFTVNLG
B_NL_3202A   MGARASVLSG .GELDKWEKI RLRPGGKKKY RLKHLVWASR ELERFAVNPG
B_TW_TWCYS   MGARASILSA .GELDKWERV RLRPGGKKRY KLKHIVWASR ELERFAVNPG
B_US_BC_L0   MGARASVLSG .GKLDKWEKI RLRPGGKKKY RLKHLVWASR ELERFAVNPG
B_US_DH123   MGARASVLSG .GKLDSWEKI RLRPGGKKKY KLKHLVWASR ELERFAVNPG
B_US_JRCSF   MGARASVLSG .GELDRWEKI RLRPGGKKKY KLKHIVWASR ELERFAVNPG
B_US_MNCG_   MGARASVLSG .GELDRWENI RLRPGGKKKY RLKHIVWASR ELERFAVNPG
B_US_P896_   MGARASVLSG .GELDRWEKI RLRPGGKKKY KLKHVVWASR ELERFAVNPG
B_US_RF_M1   MGARASVLSG .GKLDKWEKI RLRPGGKKKY KLKHIVWASR ELERFAVNPS
B_US_SF2_K   MGARASVLSG .GELDKWEKI RLRPRGKKRY KLKHIVWASR ELERFAVNPS
B_US_WEAU1   MGARASVLSG .GELDRWEKI RLRPGGKKKY KLKHIVWASR ELERFAVNPG
B_US_WR27_   MGARASMLSG .GELDKWEKI RLRPGGNKKY KLKHIVWASR ELERFAVNPG
B_US_YU2_M   MGARASVLSA .GELDKWEKI RLRPGGKKKY RLKHLVWASR ELERFAVNPG
BF1_BR_93B   MGARASVISG .GELDKWEKI RLRPGGKKQY RLKHIVWASR ELERFAVDPG
C_BR_92BR0   MGARASILRG .GKLDAWERI RLRPGGHKKY RLKHIVWASR ELERFAVNPG
C_BW_96BW0   MGARASILRG .GKLDAWERI KLKPGGKKHY MMKHLVWASR ELERFALDPG
C_BW_96BW1   MGARASILRG .GKLDAWEKI RLRPGGKKQY RIKHLVWASR ELERFALNPG
C_BW_96BW1   MGARASILRG .GKLDTWEKI RLRPGGKKRY MIKHLVWASR ELERFALNPG
C_BW_96BW1   MGARASILRG .EKLDTWEKI RLRPGGKKRY MMKHLVWASR ELERFALNPG
C_BW_96BW1   MGARASISRG .EKLDTWEKI RSRPGGKKCY MLKHIVWASR ELERFALNPA
C_ET_ETH22   MGARASILRG .EKLDAWEKI KLRPGGKKHY MLKHLVWANR ELEKFALNPD
C_IN_93IN1   MGARASILRG .GKLDKWERI RLRPGGKKHY MLKHLVWASR ELERFALNPG
C_IN_93IN9   MGARASILRG .EKLDKWEKI RLRPGGKKHY MLKHLVWASR ELDRFALNPG
C_IN_93IN9   MGARASILRG .EKLDKWERI RLRPGGKKHY MLKHLVWASR ELDRFALNPG
C_IN_94IN1   MGARASILRG .GKLDKWEKI RLRPGGKKHY MLKHLVWASR ELEKFALNPG
C_IN_95IN2   MGARASILRG .GKLDKWEKI RLRPGGKKRY MLKHLVWASR ELDRFAVNPG
CRF01_AE_C   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RMKHLVWASR ELERFALNPG
CRF01_AE_C   MGARASILSG .GKLDAWEKI RLRPGGKKQY RMKHLVWASR ELERFALNPG
CRF01_AE_C   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPG
CRF01_AE_T   MSARASVLSG .EKLNAWEKI RLRPGGRKKY KLKHLVWASR ELEKFALNPG
CRF01_AE_T   MGARASVLSG ..KLDALEKI RLRPGGKKKY KMKHLVWASR ELERFALNPG
CRF01_AE_T   MGARASVLSG .GKLDAWEKI RLRPGGKKKY KMKHLVWASR ELERFALNPG
CRF01_AE_T   MGAKASVLSG .GKLDAWEKI RLRPGGRKKY HLKHIVWASR ELERFALNPG
CRF01_AE_T   MGARASVLSG .GKLDAWEKI RLRPGGRKKY RLKHLVWASR ELERFALNPS
CRF01_AE_T   MGARASILSG .GKLDAWEKI RLRPGGRKKY RMKHLVWASR ELERFALNPG
CRF02_AG_F   MGARASVLSG .GKLDSWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPG
CRF02_AG_F   MGARASVLSG .GKLDSWEKI RLRPAGKKKY RLKHLVWASR ELERFALNPG
CRF02_AG_G   MGARVSVLRG .GQLDTWEKI RLRPGGKKKY KMKLLVWASR ELERFAVNPG
CRF02_AG_N   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPG
CRF02_AG_S   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPG
CRF02_AG_S   .GARASVLSG .GKLDAWERI RLRPGGKKKY RLKHLVWASR ELERFALNPG
CRF03_AB_R   MGARASVLSG .GKLDAWEKI RLRPGGKEKY RIKHLVWASR ELERFALNPS
CRF03_AB_R   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RIKHLVWASR ELERFAINPS
```

```
CRF04_cpx_   MGARASVLSG .GKLDAWERI RLRPGGKKKY RLKHLVWASR ELERFALNPG
CRF04_cpx_   MGARASVLSG .GKLDAWERI RLRPGGKKKY RLKHLVWASR ELERFALNPG
CRF04_cpx_   MGARASVLSG .GRLDAWEKI RLRPGGKKRY RIKHLIWASR ELERFALNPG
CRF05_DF_B   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPG
CRF05_DF_B   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHIVWASR ELERFAINPG
CRF06_cpx_   MGARASVLSG .GKLDEWEKI RLRPGGKKKY KMKHLVWASR ELERFAINPG
CRF06_cpx_   MGARASVLSG .GKLDEWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPS
CRF06_cpx_   MGARASVLSG .GKLDEWEKI RLRPGGKKKY RMKHLVWASR ELERFALNPG
CRF06_cpx_   MGARASVLSG .GKLDEWEKI RLRPGGKKKY RLKHLVWASR ELDRFALNPG
CRF11_cpx_   MGARASVLSG .GKLDSWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPS
CRF11_cpx_   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPS
D_CD_84ZR0   MGARASVLSG .GKLDAWEKI RLRPGGKKKY KLKHIVWASR ELERFALNPG
D_CD_ELI_K   MGARASVLSG .GKLDKWEKI RLRPGGKKKY RLKHIVWASR ELERYALNPG
D_CD_NDK_M   MGARASVLSG .GKLDTWERI RLRPGGKKKY ALKHLIWASR ELERFTLNPG
D_UG_94UG1   MGARASVLSG .GKLDEWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPG
F1_BE_VI85   MGARASILSG .GKLDEWEKI QLRPGGKKRY KMKHLIWASR ELERFALDPG
F1_BR_93BR   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALDPG
F1_FI_FIN9   MGARASVLSG .GKLDAWEKI RLRPGGKKQY RIKHLVWASR ELERFAIDPG
F1_FR_MP41   MGARASVLSG .GKLDAWERI RLRPGGKKKY RMKHLVWASR ELERFAVDPG
F2_CM_MP25   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHIVWASR ELKRFALNPG
F2KU_BE_VI   MGARASVLSG .GKLDSWEKI RLRPGGRKKY RLKHLVWASR ELERFALNPG
G_BE_DRCBL   MGARASVLSG .GKLDAWEKI RLRPGGKKRY RMKHLVWASR ELDRFALNPG
G_NG_92NG0   VGARASVLSG .GKLDSWEKI RLRPGGRKKY KLKHIVWASR ELGRFALNRD
G_SE_SE616   MGARASVLTG .GKLDAWEKI RLRPGGRKSY KIKHLVWASR ELERFALNPD
H_BE_VI991   MGARASVLSG .GKLDAWEKI RLRPGGRKKY RLKHLVWASR ELERFALNPD
H_BE_VI997   MGARASVLSG .GRLDTLEKI RLRPGGKKKY RLKHIVWASR ELERFALNPG
H_CF_90CF0   MGARASVLSG .GKLDAWEKI RLRPGGKKKY RLKHLVWASR ELERFALNPG
J_SE_SE702   MGARASILSG .GKLDDWEKI RLRPGGKKQY RIKHLVWASR ELDRFALNPG
J_SE_SE788   MGARASILSG .GKLDDWEKI RLRPGGKKKY RIKHLVWASR ELDRFALNPG
K_CD_EQTB1   MGARASVLSG .GKLDKWEKI QLRPGGKKKY RLKHLVWASR ELERFALNPN
K_CM_MP535   MGARASVLSG .GKLDAWEKI RLRPGGKKKY KLKHLVWASR ELERFALNPG
N_CM_YBF30   MGARASVLTG .GKLDQWESI YLRPGGKKKY RMKHLVWASR ELERFACNPG
O_CM_ANT70   MGASASVLTG .SKLDAWEQI RLKPGSKKKY RLKHLVWASR ELERFACNPE
O_CM_MVP51   MGARASVLTG .SKLDAWERI RLRPGSKKAY RLKHLVWASR ELERYACNPG
O_SN_99SE_   MGARASVLTG .SKLDAWEQI RLKPGCKKKY RLKHLVWASR ELDRSACNPE
O_SN_99SE_   MGARASVLSG .SKLDTWEQI RLKPGCKKKY RLKHLVWASR ELERFACNPE
U_CD___83C   MGARASVLSG .GKLDAWEKI RLRPGGRKKY RLKHLVWASR ELEKFAINPG

             51                                                  100
00BW0762_1   LLETSEGCKQ IIKQLQPALQ TGTEELRSLY NTVATLYCVH KKIDVRDTKE
00BW0768_2   LLETSEGCKQ IIKQLQPALQ TGTEELRSLF NTVATPYCVH EKIEVRDTKE
00BW0874_2   LLETAEGCRQ IIKQLHPALQ TGTEELRSLY NTVATLYCVH RGIEIRDTKE
00BW1471_2   LLETADGCKQ IIKQLQPALQ TGTEELRSLF NTVATLYCVH KGIKVQDTKE
00BW1616_2   LLETSDGCKQ IMKQLQPALQ TGTEELKSLF NTVATLYCVH ANIDVRDTKE
00BW1686_8   LLETSDGCKQ IIQQLQPALK TGTEELRSLF NTVATLYCAH KGISIQDTKE
00BW1759_3   LLETAEGCKR IIKQLQPALQ TGTEELKSLH NTVATLYCVH KEIDVRDTKE
00BW1773_2   LLETAEGCKQ IIKQLQPALQ TGTEELKSLY NTVATLYCVH AGIEVRDTKE
00BW1783_5   LLETSEGCKQ IIQQLQPALK TGTEELRSLY NTVATLYCVH AKIEVRDTKE
00BW1795_6   LLETSEGCKQ IMKQLQPAIQ TGTEELRSLF NTVATLYCVH EGIDVQDTKE
00BW1811_3   LLETAAGCKQ IIRQLHPALQ TGTEELRSLY NTVATLYCVH AEIEVRDTKE
00BW1859_5   LLETSEGCRQ IMRQLQPALQ TGTEELRSLF NTVATLYCVH EKIPVRDTKE
00BW1880_2   LLETAEGCKQ IIRQLHPALQ TGTQELRSLY NTVATLYCVH KDIEVRDTKE
00BW1921_1   LLETSEGCKQ IIQQLQPALQ TGTEELRSLY NTVATLYCVH KGIEVQDTKE
00BW2036_1   LLETSEGCKQ IMKQLQPALQ TGTEELRSLF NTVATLYCVH KKIQVQDTKE
00BW2063_6   LLETADGCKQ IMKQLQPALQ TGTEELRSLY NTVATLYCVH EGIDVRDTKE
00BW2087_2   LLETSEGCKQ IIKQLQPALQ TGTEELRSLF NTVATLYCVH KGIEVRDTKE
00BW2127_2   LLETSEGCRQ IIKQLQPALQ TGTEELRSLY NTIAVLYCAH QRIKVQDTKE
00BW2128_3   LLETSEGCKQ IIKQLHPALQ TGTEELKSLY NTVAVLYCVH AGIAVRDTKE
00BW2276_7   LLETSEGCKQ IMQQLQPALK TGTEELISLY NTVATLYCVH EKIQVQDTKE
00BW3819_3   LLETSEGCKQ IIKQLQPALQ TGTEELRSLY NTVATLYCVH QKIEIRDTKE
00BW3842_8   LMETSEGCKQ IIQQLQPALQ TGTEELRSLF NTIATLYCVH KGIKVQDTKE
```

```
00BW3871_3  LLETSDGCKQ IIKQLQPALQ TGTEELRSLF NTVATLYCVH KGIKVQDTKE
00BW3876_9  LLETSEGCEQ IMKQLQPALQ TGTEELRSLY NTVATLYCVH ARIEVKDTKE
00BW3886_8  LLETAEGCTQ IMKQLQPALQ TGTEELRSLF NTVATLYCVH AGIDVRDTKE
00BW3891_6  LLETSDGCKQ IIQQLQPALK TGTEELRSLY NTVATLYCVH NNIEIRDTKE
00BW3970_2  LLETSEGCKQ IIKQLQPALQ TGTEELKSLF NTVATLYCVH EKIEIRDTKE
00BW5031_1  LLETADGCKQ IIKQLQPALK TGTEELRSLY NTVATLYCVH RNIEVQDTKE
 96BW01B21  LLETSEGCKQ IMEQLQPALQ TGTEELKSLY NTVATLYCVH ENIEVRDTKE
  96BW0407  LLETSEGCKQ IIRQLHPALQ TGSEELKSLF NTVATLYCVH EKIEVRDTKE
  96BW0502  LLETAEGCKQ IMKQLQPALQ TGTEELRSLY NTVATLYCVH AGIEVRDTKE
 96BW06_J4  LLETSEGCKQ IMQQLQPALQ TGTEELRSLF NTVATLYCVH KGIKVQDTKE
 96BW11_06  LLETSEGCKQ IMKQLQPALQ TGTEEIRSLY NTVATLYCVH EGIEVQDTKE
  96BW1210  LLETAEGCKQ IMKQLQPALK TGTEELKSLY NTVATLYCVH EKIEVRDTKE
 96BW15B03  LLETSEGCKQ IMKQLQPALQ TGTEELRSLF NTVATLYCVH EKIKVKDTKE
 96BW16_26  LLETSEGCRQ IMIQIQPALK TGTEELKSLY NTVAVLFCVH AGIEVQDTKE
 96BW17A09  LLETAEGCKQ IIKQLQPALQ TGTEELKSLY NTVATLYCVH QGIEVRDTKE
 96BWMO1_5  LLETSEGCKQ IIKQLQPTIQ TGTEELKSLY NTVATLYCVH KGIDVRDTKE
 96BWMO3_2  LLETSEGCKQ IIKQLQPALQ TGTEELRSLY NTVATLYCVH EKIEVRDTKE
98BWMC12_2  LLETAEGCTQ IMRQLQPALQ TGTEELRSLY NTVATLYCVL KGIKVRDTKE
98BWMC13_4  LLETSEGCKQ IMRQLQPAIQ TGTEELRSLF NTVATLYCVH EGIDVKDTKE
98BWMC14_a  LLETSEGCQQ IMKQLQPALQ TGTEELRSLF NTVATLYCVH MGIEVRDTKE
98BWMO14_1  LLETAEGCKQ IIQQLQPALK TGTEELRSLF NTVATLYWVH KNVEVRDTKE
98BWMO18_d  LLETSEGCRK IMKQLQPSIQ TGTEELRSLY NTIATLYCVH EKIEVRDTKE
98BWMO36_a  LLETAEGCKQ ILQQLQPALK TGTEELKSLY NTVATLYCVH QGIEVRDTKE
98BWMO37_d  LLENSEGCKQ IMKQLQPALQ TGTEELKSLF NTVATLYCVH EKIEVRDTKE
99BW3932_1  LLETADGCKQ IIKQLQPAFQ TGTEEIRSLY NSIAVLYCVH RRLTIQDTKE
99BW4642_4  LLETSEGCKQ IMKQLQPALQ TGTEELRSLY NTVATLYCVH ERIQVRDTKE
99BW4745_8  LLETTEGCRQ IIRQLQPALQ TGTEELKSLY NTVATLYCVH NNIEIRDTKE
99BW4754_7  LLETAEGCKQ IIKQLHPALQ TGTEELRSLY NTVATLYCVH NKIEVQDTKE
99BWMC16_8  LLETSEGCKQ IMKQLQPALQ TGTEELRSLY NTIATLYCVH ESIEVRDTKE
A2_CD_97CD  LLETETGCRR IFGQLQPALE TGTEELRSLY NTIAVLYFVH QKIEVKDTKE
A2_CY_94CY  LLETPEGCRQ IIRQLQPALQ TGTEELKSLY NTVVVLYWVH QRVDVKDTKE
A2D___97KR  LLETSEGCKQ IIGQLEPSIK TGSEEIKSLF NAVVTLYCVH QRIEVRDTKE
A2G_CD_97C  LLETAAGCRQ ILGQLQPALQ TGTEELRSLY NTVAVLYCVH QKIEVKDTKE
A_BY_97BL0  LLETSEGCQQ ILEQLQPTLK TGSEELKSLY NTVATLYCVH QRIEIKDTKE
A_KE_Q23_A  LLETADGCQQ IMEKLQPALK TGTEEIKSLF NTVATLYCVH QRIDVKDTKE
A_SE_SE659  LLETTEGCQQ IIEQLQPALN TGTEEIKSLF NTVATLYCVH QRIDVKDTKE
A_SE_SE725  LLETTEGCQQ IMEQLQSALK TGTEELRSLF NTVAVLYCVH QRIEIKDTKE
A_SE_SE753  LLETAEGCQQ IMKQLQPALK TGTEELRSLY NTVATLYCVH QRINVTDTKE
A_SE_SE853  LLETTEGCQQ IIGQLQPAFK TGTEELKSLY NTVATLWCVH QRIDVKDTKE
A_SE_SE889  LLETTEGCQQ IMDKLQPALR TGTEELRSLY NTVATLYCVH QKIAVKDTKE
A_SE_UGSE8  LLETAEGCQQ IIEQLQSAPK TGTEELRSLY NTVATLYCVH QRIDVKDTKE
A_UG_92UG0  LLETTEGCQQ IMEQLQSALR TGTEELRSLY NTVATLYCVH QRIEVKDTKE
A_UG_U455_  LLETAEGCQQ ILGQLQPALQ TGTEELRSLY NTVAVLYCVH QRIDVKDTKE
AC_IN_2130  LLETSEGCKQ IIKQLQPALQ TGTEELRSLH NTVATLYCVH AGIEIRDTKE
AC_RW_92RW  LLETPEGCKQ IMRQLQPALQ TGTDELRSLY NTVATLYCVH QKIDVKDTKE
AC_SE_SE94  LLETSEGCQQ ILEQLQPALK TGTEEIKSLF NTVATLYCVH QRIEVKDTKE
ACD_SE_SE8  LLETTEGCQQ IMDQLQPALK TGTEELRSLY NTVATLWCVH KRIDVKDTKE
ACG_BE_VI1  LLETAEGCQQ IIEQLQSTLK TGSEELKSLF NTVATLWCVH LRIEIKDTKE
AD_SE_SE69  LLETPEGCKQ IMGQLQPAIQ TGSEELKSLF NTVATLYCVH AKIKVTDTKE
AD_SE_SE71  LLETTKGCQQ IMEQLQPAFK TGTEELKSLY NTVATLYCVH QQIEVKDTKE
ADHK_NO_97  LLETSEGCQQ VMEQLQPALK TGSEELKSLF NLVAVLWCVH QRIDVRDTKE
ADK_CD_MAL  LLETGEGCQQ IMEQLQSTLK TGSEEIKSLY NTVATLYCVH QRIDVKDTKE
AG_BE_VI11  LLETAEGCQQ IIEQLQSTLR TGSEELKSLY NTLATLWCVH KKIEVRDTKE
AG_NG_92NG  LLETTEGCQQ IMRQLQPSLQ TGTEEIKSLF NTVATLYCVH QRIEVKDTKE
AGHU_GA_VI  LLETAEGCQQ IMEQLQSALK TGSEELKSLF NTVATLYCVH QKIDVKDTKE
AGU_CD_Z32  LLETTEGCQQ ILSQFQPALK TGTEEIISLY NTVATLYCVH QNIEVKDTKE
AJ_BW_BW21  LLETAEGCQQ IMEQLQSALK TGSEELRSLY NTVATLYCVH QRIEVKDTKE
B_AU_VH_AF  LLETSEGCRQ ILVQLQPSLP TGSEELKSLF NTVATLYCVH QKIEIRDTKE
B_CN_RL42_  LLETSEGCRQ ILEQLQPALQ TGSEELRSLF NTVATLYCVH LRIEIKDTKE
B_DE_D31_U  LLETSEGCRQ ILGQLQPSLQ TGSEELRSLF NTVATLYCVH QRIEVKDTKE
B_DE_HAN_U  LLETSEGCRQ IMGQLQPSLQ TGSEELRSLY NTVATLYCVH QKIEVKDTKE
```

```
B_FR_HXB2_  LLETSEGCRQ ILGQLQPSLQ TGSEELRSLY NTVATLYCVH QRIEIKDTKE
B_GA_OYI__  LLETSEGCRQ ILGQLQPSLK TGSEEIRSLY NTVATLYCVH QKIEVKDTKE
B_GB_CAM1_  LLETSEGCRQ ILGQLQPSLQ TGSEELRSLY NTVATLYCVH QKIDVKDTKE
B_GB_GB8_A  LLETSEGCRQ ILEQLQPSLQ TGSEERRSLF NTVATLYCVH QRIDVKDTKE
B_GB_MANC_  LLETSEGCRQ ILVQLQPSLQ TGSEELKSLY NTVAVLYCVH QGIEVKDTKE
B_KR_WK_AF  LLETSEGCRQ ILGQLQPSLQ TGSEELKSLF NAVAVLYCVH QGIEVKDTKE
B_NL_3202A  LLETAEGCRQ ILGQLQPALQ TGSEELRSLF NTVAVLYCVH QRIEIKDTKE
B_TW_TWCYS  LLETSEGCRQ ILGQLQSSLQ TGSEELRSLY NTVATLYCIH QKIEIKDTKD
B_US_BC_L0  LLETAEGCRQ ILGQLQPSLQ TGSEELKSLY NTIAVLYCVH QMIDVKDTKE
B_US_DH123  LLETSEGCRQ ILGQLQPSLQ TGSEELRSLY NTVATLYCVH ERIEVKDTKE
B_US_JRCSF  LLESSEGCRQ ILGQLQPSLK TGSEELTSLY NTVATLYCVH QRIEIKDTKE
B_US_MNCG_  LLETSEGCRQ ILGQLQPSLQ TGSEELKSLY NTVATLYCVH QKIEIKDTKE
B_US_P896_  LLETSEGCRQ ILGQLQSSLQ TGSEELKSLY NTVATLYCVH QRIEVKDTKE
B_US_RF_M1  LLETAEGCRQ ILGQLQPALQ TGSEELKSLY NAVATLYCVH QNIEVRDTKE
B_US_SF2_K  LLETSEGCRQ ILGQLQPSLQ TGSEELRSLY NTVATLYCVH QRIDVKDTKE
B_US_WEAU1  LLETSEGCRQ ILGQLQPSLQ TGSEELRSLY NTVAVLYCVH QKIEVKDTKE
B_US_WR27_  LLETAEGCRQ ILRQLQPSLQ TGSEELRSLF NTVATLYCVH QRIGVKDTKE
B_US_YU2_M  LLETSEGCRQ ILGQLQPSLQ TGSEELRSLY NTVATLYCVH QKIEVKDTKE
BF1_BR_93B  LLETSEGCRQ ILEQLQPALK TGSEELRSLY NTVATLYCVH QKIDVKDTKE
C_BR_92BR0  LLETSEGCKQ IMKQLQPALQ TGTKELISLH NTVATLYCVH EKIDVRDTKE
C_BW_96BW0  LLETSEGCKQ IIQQLQPALQ TGTEELRSLF NTVATLYCVH KKIEVRDTKE
C_BW_96BW1  LLGTSEGCKQ IMKQLQPALQ TGTEEIRSLY NTVATLYCVH EGIEVQDTKE
C_BW_96BW1  LLETAEGCKQ IMKQLQPALK TGTEELKSLY NTVATLYCVH EKIEVRDTKE
C_BW_96BW1  LLETSEGCKQ IMKQLQPALQ TGTEELRSLF NTVATLYCVH EKIKVKDTKE
C_ET_ETH22  LLDTSAGCKQ IIKQLQPALQ TGTEELKSLF NTVATLYCVH QKIEIKDTKE
C_IN_93IN1  LLETAEGCKQ IIKQLQPALQ TGTEELKSLH NTVATLYCVH AGIEVRDTKE
C_IN_93IN9  LLETSEGCKQ IIKQLQPALQ TGTEELRSLH NTVVTLYCVH AGIEVRDTKE
C_IN_93IN9  LLEAAEGCKQ IIKQLQPALQ TGTEELRSLH NTVATLYCVH EGIEVRDTKE
C_IN_94IN1  LLETSEGCKQ IMKQLQPALQ TGTEELRSLF NTVATLYCVH AEIEVRDTKE
C_IN_95IN2  LLETAEGCKQ IIKQLQPALQ TGTEELRSLF NTVATLYCVH AGIEVRDTKE
CRF01_AE_C  LLETAEGCQQ LIEQLQSNLK TGSEEIKSLF NTIATLWCVH QRIDVKDTKE
CRF01_AE_C  LLETAEGCQQ LIEQLQSTIK TGSEELKSLF NTIATLWCVH QRIDVKDTKE
CRF01_AE_C  LLETAEGCQQ LLEQLQSTLK TGSEELKSLF NTIATLWCVH QRINVTDTKE
CRF01_AE_T  LLETAEGCQQ LIEQLQSTLK TGSEELKSLF NTIATLWCVH QKIEVKDTKK
CRF01_AE_T  LLETAEGCQQ IIEQLQSALK TGSEELKSLY NTVVTLWCVH QRIDVKDTKE
CRF01_AE_T  LLETAEGCQQ LIEQLQSTLK TGSEELKSLY NTIATLWCVH QRIEVKDTKE
CRF01_AE_T  LLETAEGCQQ LIEQLQSTLK TGSEELRSLF NTIVTLWCVH QRIEVKDTKE
CRF01_AE_T  FLETAEGCQQ IIEQLQSTLK TGLEELKSLF NTVATLWCVH QRIEVKDTKE
CRF01_AE_T  LLETAEGCQQ LIGQLQSNLK TGSEELRSLF NTVATLWCVH QRIEVKDTKE
CRF02_AG_F  LLETAEGCQQ LIEQLQTALG TGSEELKSLY NTIATLWCVH KRIDIKDTKE
CRF02_AG_F  LLETAEGCQQ LMKQLQSALG TGSEELKSLY NTIATLWCVH QRIDIKDTKE
CRF02_AG_G  LLETTEGCQQ ILEQLQPALK AGSEELKSLY NTVATLYCVH QKIDVRDTKE
CRF02_AG_N  LLETAEGCQQ LMEQLQSALR TGSEELKSLF NTIATLWCVH QRIDIKDTKE
CRF02_AG_S  LLETAEGCQQ LMEQLQSTLR TGSEELKSLF NTIATLWCVH QRIDIKDTKE
CRF02_AG_S  LLETAEGCQQ IIEQLQSTLK TGSEELKSLF NTVATLYCVH QRIEVKDTKE
CRF03_AB_R  LLETSEGCQQ ILEQLQPTLK TGSEELKSLY NTVATLYCVH QRIEIKDTKE
CRF03_AB_R  LLETSEGCQQ ILEQLQPTLK TGSEELKSLY NTVATLYCVH RRIEIKDTKE
CRF04_cpx_  LLETAEGCQQ LMEQLQSTLK TGSEELRSLY NTITTLWCVH QRIDVQDTKE
CRF04_cpx_  LLETAEGCQQ LMEQLQPTLR TGSEEVKSLF NTLATLWCVH QRIDVKDTKE
CRF04_cpx_  LLETAEGCQQ LMEQLQSTLK TGSEELKSLF NLIATLWCVH QRIEVKDTKE
CRF05_DF_B  LLETSEGCKQ IISQLQPAIQ TGTEELKSLY NTVATLYCVH ERIEVKDTKE
CRF05_DF_B  LLETSEGCKQ IIAQLQPAIQ TGSEELRSLY NTVATLYCVH EKIEVKDTKE
CRF06_cpx_  LLETAEGCQQ IIEQLQSTLK TGSEELKSLF NTVATLYCVH QRIKVTDTKE
CRF06_cpx_  LLETAEGCQQ IMEQLQSALR TGSEELKSLY NTVATLYCVH QRIKVTDTKE
CRF06_cpx_  LLETPEGCQQ IMEQLQSTLR TGSEELKSLY NTVATLYCVH QRIKVTDTKE
CRF06_cpx_  LLETAEGCQQ IIEQLQSALK TGSEELRSLF NTVATLYCVH QRIKVQDTKE
CRF11_cpx_  LLETGEGCQQ LMGQLQPALG TGTEELRSLY NTLATLYCVH HRIEIKDTKE
CRF11_cpx_  LLETGEGCKQ IMGQLQPALG TGTEELRSLY NTVAALYCVH LRIEVKDTKE
D_CD_84ZR0  LLETSEGCKQ IMEQLQPALK TGSEELRSIY NTVATLYCVH KRIEIKDTKE
D_CD_ELI_K  LLETSEGCKQ IIGQLQPAIQ TGTEELRSLY NTVATLYCVH KGIDVKDTKE
D_CD_NDK_M  LLETSEGCKQ IIGQLQPSIQ TGSEEIRSLY NTVATLYCVH ERIEVKDTKE
```

```
D_UG_94UG1   LLETSEGCRQ IIRQLQPSIQ TGSEEIKSLY NTVVTLYCVH ERIKVASTKE
F1_BE_VI85   LLETSEGCQK IIRQLQPSLQ TGSEELKSLF NTVAVLYYVH QRAGVTDTKE
F1_BR_93BR   LLETSEGCRK IIGQLQPSLQ TGSEELKSLY NTIAVLYYVH QKVEVKDTKE
F1_FI_FIN9   LLETSEGCQK IIAQIQPSIQ TGSEELRSLY NTIAVLYFVH QKIEVKDTKE
F1_FR_MP41   LLETPEGCKQ IIRQLQPSLQ TGSEELRSLF NTVAVLYCVH QKIEIKDTKE
F2_CM_MP25   LLETTEGCKK IIGQLQPSLQ TGSEELKSLF NTIVVLYYVH QKIEVRDTKE
F2KU_BE_VI   LLESANGCRQ IIEQIQPALQ TGSEELRSLF NTVVTLYWVH QRIEVKDTKE
G_BE_DRCBL   LLETAEGCQK IMAQLQPALQ TGTEEIKSLF NTVATLYCVH QKIEVRDTKE
G_NG_92NG0   LLETAEGCVQ IMKQLQP.AL TGTEELRSLF NTVATLYCVH QKIEVKDTKE
G_SE_SE616   LLETAEGCQQ IMRQLQPSLQ TGTEEIKSLY NAVATLYCVH QRIEVKDTKE
H_BE_VI991   LLETADGCQQ ILGQLQPALK TGTEDLQSLY NTIAVLYCVH QRIDVKDTKE
H_BE_VI997   LLESAEGCLQ IIEQLRPSIK TGTEELXSLF NTVATLYCVL QRIEVKDTKE
H_CF_90CF0   LLETPEGCLQ IIEQIQPAIK TGTEELKSLF NLVAVLYCVH RKIDVKDTKE
J_SE_SE702   LLESAKGCQQ ILVQLQPALQ TGTEEIKSLY NTVATLYCVH QRIEIKDTKE
J_SE_SE788   LLESAKGCQQ ILVQLQPALQ TGTQEIKSLY NTVATLYCVH QRIEIKDTME
K_CD_EQTB1   LLETVEGCRQ IIRQLQPSLQ TGSEELRSLF NTVATLYWVH QSIQVRDTKE
K_CM_MP535   LLETTEGCRQ IITQIQPSIQ TGSEEIKSLY NTIAVLYFVH QKIEVKDTKE
N_CM_YBF30   LMDTADGCAK LLNQLEPALK TGSEELRSLY NALAVLYCVH SRIQIHNTQE
O_CM_ANT70   LLETAEGNEK LLQQLEPALK TGSDSLQSLW NAIVVLWCVH NRYKIGDTQQ
O_CM_MVP51   LLETAEGTEQ LLQQLEPALK TGSEDLKSLW NAIAVLWCVH NRFDIRDTQQ
O_SN_99SE_   LLETAEGNEK LLQQLEPALK TGSESLQSLW NTIAVLWCVH NRIKVEDTQQ
O_SN_99SE_   LLETAEGNEE LLQQLEPALK TGSESLQSLW NTIAVLWCVH KRFKVEDTQQ
U_CD___83C   LLETAEGCQQ IIGQLQPALQ TGSEELKSIY NLVATLYCVH QRIEVKDTKE

             101                                                  150
00BW0762_1   ALDQIEKEQN QSQQK..... .TQQAETADK K..........  VSQNYPIVQN
00BW0768_2   ALDKMEEEQN KSQQK..... TQQAEAA.AG K..........  VSQNYPIVQN
00BW0874_2   ALDKIEKEQN ......ESQQ KTQQAKAADG K..........  VSQNYPIVQN
00BW1471_2   ALDKIEEEQN K.......... CQQQAKAAEG K..........  VSQNYPIVQN
00BW1616_2   ALDKLEEEQN KSQQE..... TQQQTKAADG KIS.......  ..HNYPIVQN
00BW1686_8   ALDKIEEEQN KSQQ...... KTQQVAAAAG Q..........  VSQNYPIVQN
00BW1759_3   ALDKIEEEQK KSQQK..... .LQQEEAADR KV.......S  ..QNYPIVQN
00BW1773_2   ALDKIEEEQN KCQQK..... .TQQAKEADG K..........  VSQNYPIVQN
00BW1783_5   ALDKIEEEQN KCQQK..... .TQQAEKTQQ AGAA...GGK  VSQNYPIVQN
00BW1795_6   ALDKVEEEQN KSQQ...... KMQQAEAADK KVS.......  ..QNYPIVQN
00BW1811_3   ALDKIEEEQN KSQQK..... .TQQAQEATA TG.......K  VSQNYPIVQN
00BW1859_5   ALDKIEEEQN KCQQ...... KTQQTEAAAG K..........  VSQNYPIVQN
00BW1880_2   ALDKIEEEQN KSQQK..... .AQQAEAADK Q..........  VSQNYPIVQN
00BW1921_1   ALDKIEEEQN KSQQ...... KTQQAEAAAG K..........  VSQNYPIVQN
00BW2036_1   ALDKIEEEQN TCQQ...... KTQQAEAAAG K..........  ISQNYPIVQN
00BW2063_6   ALDKIEEEQQ KSQQ...... KTQQAEAADK KVS.......  ..QNYPIVQN
00BW2087_2   ALDKIEEEQN KCQQK..... .TQQAKAADK G........E  VSQNYPIVQN
00BW2127_2   ALDEIEEEQN KCQQK..... .AQQEKATDG .........K  VSQNYPIVQN
00BW2128_3   ALDKIEEEQN KIQQK..... .TQQAKEADK .........K  VSQNYPIVQN
00BW2276_7   ALDKIEEEQN KSQQKSQQ.. KTQQAQAADG K..........  VSQNYPIVQN
00BW3819_3   ALDKIEEEQN KCQQK..... .TQQGKAADE .........K  ISQNYPIVQN
00BW3842_8   ALDKIEEEQN KSQQKTQQQQ KTQQTEAAAG K..........  VSQNYPIVQN
00BW3871_3   ALDKIEEEQN KSQQ...... KTQQAEAAAE K..........  VSQNYPIVRN
00BW3876_9   ALDKIEEEQN KRQQK..... .AQQEKAADG .........K  VSQNYPIVQN
00BW3886_8   ALDKIEEEQN KSQQK..... .TQQAQAADK .........K  VSQNYPIVQN
00BW3891_6   ALDRIEEEQK KCQQK.TQQQ KTQQVEAADG K..........  VSQNYPIVQN
00BW3970_2   ALDKIEEEQK KSQQK..... .TQQAKAADG KI.......S  ..QNYPIVQN
00BW5031_1   ALDKVEEEQN NCQQK..... .TQQAEAAGG .........K  ISQNYPIVQN
 96BW01B21   ALDKIEEEQN KSQQK..... .TQQAKTDD. .G.......K  ISPNYPIVQN
  96BW0407   ALDKIEEEQN KCQQK..... .IQQAEAADK G........K  VSQNYPIVQN
  96BW0502   ALDKIEEEQN KSQQK..... .TQQAKEADG K..........  VSQNYPIVQN
 96BW06_J4   ALDKIEEEQN KSQQ...... KTQQAEAAAG K..........  VSQNYPIVQN
 96BW11_06   AVDKIEEEQN KSQQ...... KTQQAEAADK KVS.......  ..QNYPIVQN
  96BW1210   ALDKIEEEQN KCQQK..... .EQQAEAAAK G........K  VSQNYPIVQN
 96BW15B03   ALDKIEEEQN KSQE...... KTQQAEAAAG K..........  ISQNYPIVQN
 96BW16_26   ALDKIEEEQN KSQQK..... .TQQAKEATG K..........  VSQNYPIVQN
```

```
 96BW17A09  ALDKIEEEQN K.........  FQQQAKAADG K.........  VSQNYPIVQN
 96BWMO1_5  ALDKIEEEQN KIQQ......  KMQQAEAADK KISQ...DKK  ISQNYPIVQN
 96BWMO3_2  ALDRIEEEQN KCQQK.....  AQQQEEA.AA K.........  VSQNYPIVQN
98BWMC12_2  ALDKIKEEQN ......KSKQ  KTQQAKAADE K.........  VSQNYPIVQN
98BWMC13_4  ALDKVEEEQN KIQQ......  KMQQAESADK KVS.......  ..QNYPIVQN
98BWMC14_a  ALDKIEEEQN KVQQ......  KTQQAEAAAG K.........  VSQNYPIVQN
98BWMO14_1  ALDKIEEEQN KSQQK.....  TQQAEAADKG K.........  VSQNYPIVQN
98BWMO18_d  ALDKVEEEQK QCQQR.....  .TQQAEAADK G........K  VSQNYPIVQN
98BWMO36_a  ALDKIEEEQN KIQQQ.....  .TQKATAADK G........K  VSQNYPIVQN
98BWMO37_d  ALDKIEEEQN KSQQ......  ..K.TQQADG KV.......S  ..QNYPIVQN
99BW3932_1  ALDKIEEEQN KCQQK.....  .TQQEKMAEG .........K  VSQNYPIVQN
99BW4642_4  ALDKIEEEQN KCQS......  KTQQAEAAAE K.........  VSQNYPIVQN
99BW4745_8  ALDRIEEEQN ......KSQQ  KAQQAKAAEG K.........  VSQNYPVVQN
99BW4754_7  ALDKIEEEQN KSQQK.....  .TTQAEAADK K.........  VSQNYPIVQN
99BWMC16_8  ALDKIEEEQN KCQQK.....  .TQQAEAADK .G.......K  VSQNYPIVQN
A2_CD_97CD  ALDKIEEEQN KCKQK.....  .TQQAAADTG SSSSQNYRGS  SSQNYPIVQN
A2_CY_94CY  ALDKIEEEQN ..KQK.....  .TQHAAADTG NS.......S  S.QNYPIVQN
A2D___97KR  ALDKLEEEQN KHKQK.....  .TQPAAADTG SS........  GSQNYPIVQN
A2G_CD_97C  ALDKIEEEQN TCKQR.....  .TQHAAADTG SSRSQDYRGS  SSQNYPIVQN
A_BY_97BL0  ALDKIEEIQN XSKQ......  KTQQAATGTG S.....SS.K  VSQNYPIXQN
A_KE_Q23_A  ALDKIEEIKN KSKQ......  KTQQAAADTG N.....SS.N  VSQNYPIVQN
A_SE_SE659  ALDKIEEMQK KSKQ......  KTSQATADTG S.....SS.K  VSQNYPIVQN
A_SE_SE725  ALDKIEEIQK KSKQ......  KAQQAAADTG N.....SS.K  VSQNYPIVQN
A_SE_SE753  ALDKIEEIQN KSEQ......  KTQQAAADTG N.....S..K  VSQNYPIVQN
A_SE_SE853  ALDKLEEIQK KSKQ......  KTQQAVADTG S.....SS.K  VSHNYPVVQN
A_SE_SE889  ALDKIEEIQN KNKQ......  KTQQAAADTG N.....SS.N  VSRNYPIVQN
A_SE_UGSE8  ALDKIEEIQN KSKQ......  KKEQAAADTG N.....SS.K  VSQNYPIMQN
A_UG_92UG0  ALDKIEEIQK KSKQ......  KTQQAAADTG S.....SS.K  VSQNYPIVQN
A_UG_U455_  ALNKIEEMQN KNKQR.....  .TQQAAANTG SS........  ..QNYPIVQN
AC_IN_2130  ALDKIEEEQE KSQQK.....  .TQQAKEADG K.........  VSQNYPIVQN
AC_RW_92RW  ALDKIEEEQN KSQQK.....  .TQQAEAADK G........K  VSQNYPIVQN
AC_SE_SE94  ALDKIEEIKN KSKQ......  KAQQAAADTG N.....SG.K  VSQNYPIVQN
ACD_SE_SE8  ALDKIEEIQK KSKQ......  KAQQAAADTG N.....SS.N  VSQNYPIVQN
ACG_BE_VI1  ALDKLEEVQN KSKQR.....  .TQQAAAATG .........S  GSQNFPIVQN
AD_SE_SE69  ALDKIEEEQT KSK......K  KAQQATADTK N.....SS.Q  VSQNYPIVQN
AD_SE_SE71  ALDKIEEMQN RSKQ......  KTQQAAADTG N.....NS.K  VSQNYPIVQN
ADHK_NO_97  ALDKIEEIQN KSKQ......  KTQQAAAATG N.....GS.N  ISQNYPIVQN
ADK_CD_MAL  ALDKIEEIQN KSRQKTQQAA  AAQQAAAATK N.....SS.S  VSQNYPIVQN
AG_BE_VI11  AVDKIEEMQN KSKQ......  KTQQAAAADT G........S  SSQNYPIVQN
AG_NG_92NG  ALEEVEKIQK NSQQE.....  .TQKAAMGKG N.....SS.Q  VSQNYPIVQN
AGHU_GA_VI  ALDKIEEIQN KSKQ......  KTQQAAAATG N.....SS.K  VSQNYPIVQN
AGU_CD_Z32  ALDEVEKAQK SSQQK.....  .TQKATMAEE N.....SSNQ  VSKNFPIVQN
AJ_BW_BW21  ALDKIEEIQN KSKQ......  KAQQAAAATG N.....SS.S  VSQNYPIVQN
B_AU_VH_AF  ALEKIEEEQN KSKKKAQQ.A  AADAAAADAG N.....SS.K  VSQNYPIVQN
B_CN_RL42_  ALEKIEEEQN KSKK......  KAQQTAAGTG NNS......Q  ASQNYPIVQN
B_DE_D31_U  ALDKIEEEQN KSKK......  KAQPAARDAG N.....NS.Q  VSQNYPIVQN
B_DE_HAN_U  ALDKVEEEQN NSKK......  KAQQEAADAG N.....RN.Q  VSQNYPIVQN
B_FR_HXB2_  ALDKIEEEQN KSKK......  KAQQAAADTG H.....SN.Q  VSQNYPIVQN
B_GA_OYI__  ALDKIEEEQN KSKK......  KAQQTAADTG NSS......Q  VSQNYPIVQN
B_GB_CAM1_  ALEKIEEEQN KSKK......  KAQQAAAGTG N.....SS.Q  VSQNYPIVQN
B_GB_GB8_A  ALDKIEEEQS KSKK......  KAQQAAADKG NSNS...N.Q  VSQNYPIVQN
B_GB_MANC_  ALDKIEEEQN KSKK......  KAQPAAADTG N.....SS.Q  VSHNYPIVQN
B_KR_WK_AF  ALEKIEEEQS KSKK......  KAQQATADTG S.....SS.Q  VSQNYPIVQN
B_NL_3202A  ALDKIEEEQN KSKK......  KAQQAAADTG N.....SS.Q  VSQNYPIVQN
B_TW_TWCYS  ALEKIEEEQN KSKK......  KAQQAATGTG NNS......P  VSQNYPIVQN
B_US_BC_L0  ALEKIEEEQN NSKK......  KAQQAAADTG N.....SS.Q  VSRNYPIVQN
B_US_DH123  ALDKVEEEQN KSKK......  KAQQAAADTG N.....SS.Q  VSQNYPIVQN
B_US_JRCSF  ALEKIEEEQT KSMK......  KAQQAAADTG N.....SS.Q  VSQNYPIVQN
B_US_MNCG_  ALEKIEEEQN KSKK......  KAQQAAADTG NRG.N.SS.Q  VSQNYPIVQN
B_US_P896_  ALDKIEEEQN KSKK......  KAQQAAADTG N.....SS.Q  VSQNYPIVQN
B_US_RF_M1  ALDKIEEEQN KSKK......  KAQQAAADTG N.....GS.Q  VSQNYPIVQN
```

```
B_US_SF2_K  ALEKIEEEQN KSKK..... KAQQAAAAAG TG..N.SS.Q VSQNYPIVQN
B_US_WEAU1  ALDKIEEEQN KSKK..... KAQQAAADTE N.....RN.Q VSQNYPIVQN
B_US_WR27_  ALEKIEEEQN KSKK..... KAQQAEADTG N.....NS.K VSQNYPIVQN
B_US_YU2_M  ALEKIEEEQN KSKK..... KAQQAAADTG N.....SS.Q VSQNYPIVQN
BF1_BR_93B  ALEKIEEEQN .SKK..... KAQQAAANTG NNS.....Q VSQNYPIVQN
C_BR_92BR0  ALDKIKEEQN KSQQK..... .TQQAEAADK G........K VSQNYPIVQN
C_BW_96BW0  ALDKIEEEQN KCQQK..... .IQQAEAADK G........K VSQNYPIVQN
C_BW_96BW1  AVDKIEEEQN KSQQ..... KTQQAEAADK RVS..... ..QNYPIVQN
C_BW_96BW1  ALDKIEEEQN KCQQK..... .EQQAEAAAK G........K VSQNYPIVQN
C_BW_96BW1  ALDKIEEEQN KSQE..... KTQQAEAAAG K..... ISQNYPIVQN
C_ET_ETH22  ALDKIEEEQN ESQQK..... .TQQAGAADR G........K DSQNYPIVQN
C_IN_93IN1  ALDKIEEEQN KIQQK..... .TQQAKEADG K..... VSQNYPIVQN
C_IN_93IN9  ALDKIEEEQN KNQQK..... .TQQAKEADE K..... VSQNYPIVQN
C_IN_93IN9  ALDKIEEEQN KSQQK..... .TQQAKEADG K..... VSQNYPIVQN
C_IN_94IN1  ALDKIEEEQN KSQQK..... .TQQAKEADS ..... ..QNYPIVQN
C_IN_95IN2  ALDKIEEEQN KIKQK..... .TQQAKEDDG K..... VSQNYPIVQN
CRF01_AE_C  ALDKIEEVQN KNKQ..... KTQQAAAGTG S.....NS.K VSQNYPIVQN
CRF01_AE_C  ALDKIEEVQN KSKQ..... KAQQAAAGTG S.....SS.K VSQNYPIVQN
CRF01_AE_C  ALDKIEEIQN KNKQ..... KTQQAAADTG S.....SS.K VSQNYPIVQN
CRF01_AE_T  ALDKIEEVQK KSQQ..... KTQQAAAGTG S.....SS.K VSQNYPIVQN
CRF01_AE_T  ALDKIEEVQN KSQQ..... KTQQAAAGTG S.....SS.K VSQNYPIVQN
CRF01_AE_T  ALDKIEEVQK KSQQ..... KKQQAAADTG S.....SS.K VSQNYPIVQN
CRF01_AE_T  ALDKIEEVQN KSQQ..... KTQQAAAGTG S.....SS.T VSHNYPIVQN
CRF01_AE_T  ALDKIEEVQN KSQR..... KTQQAAAGTG S.....SS.K VSQNYPIVQN
CRF01_AE_T  ALDKIEEAQK KSQQ..... KTQQAAAGTG S.....SS.K VSQNYPIVQN
CRF02_AG_F  ALEKIEEVQN KSKQ..... KTQQAAAATG .........S SSQNYPIVQN
CRF02_AG_F  ALDKIEEVQN KSKQ..... KTQQAEAATG .........S SSQNYPIVQN
CRF02_AG_G  ALDKLEEIQN KSKQK..... .TQQAAANS. .........Q VSQNYPIVQN
CRF02_AG_N  ALDKMEEIQN KSKQ..... KVQQTAAATG .........S SSQNYPIVQN
CRF02_AG_S  ALDKVEEVQK KSKQ..... KTQQAAAAAG .........S SSQNYPIVQN
CRF02_AG_S  ALDKLEEIQE KSKQ..... KTQQAAAATG .........S SSQNFPIVQN
CRF03_AB_R  ALDKIEEIQN KSKQ..... KTQQAATGTG S.....SS.K VSQNYPIVQN
CRF03_AB_R  ALDKIEEIQN KSKQ..... KTQQAATGTG S.....SS.K VSQNYPIVQN
CRF04_cpx_  ALDKIEEIQS KSKQ..... KTQQAAAAAG G.....SS.N VSQNYPIVQN
CRF04_cpx_  ALDKVEEVQK KSKQ..... KTQQAAADTG G.....SS.K VSQNYPIVQN
CRF04_cpx_  ALDKVEEMQN KSKQ..... KTQQAAAGN. .....SS.N VSQNYPIAQN
CRF05_DF_B  ALEKIEEEQN KSKR.....K KAQQAEAGAG N.....SS.Q ASQNYPIVQN
CRF05_DF_B  ALEKIEEEQN KSKS.....K RAQQAEAGTK N.....SG.P VSQNFPIVQN
CRF06_cpx_  AIDKIEEIQK KSKQK..... .AHQAAATG. N.....SS.N LSQNYPIVQN
CRF06_cpx_  ALDKIEEIQN KSKQK..... .AQQAAAATG N.....SS.N LSQNYPIVQN
CRF06_cpx_  ALHKIEEIQN KSKQK..... .VQQAAAATG N.....SS.Q LSQNYPIVQN
CRF06_cpx_  AVDKVEEIQN KSKQK..... .TQQTTAATG N.....SS.N LSQNYPIVQN
CRF11_cpx_  ALDKIEEIQN KSKQ..... KKQQAAADTG N.....SS.N VSQNYPIVQN
CRF11_cpx_  ALDKVEEIQN KSK..... .KQQAAADSG N.....SN.K VSQNYPIVQN
D_CD_84ZR0  ALDKIEEEQN NSKK..... RAQQAAAGTG N.....SS.Q VSQNYPIVQN
D_CD_ELI_K  ALEKMEEEQN KSK.....K KAQQAAADTG N.....NS.Q VSQNYPIVQN
D_CD_NDK_M  AVEKMEEEQN KSK.....K KTQQAAADS. .....S.Q VSQNYPIVQN
D_UG_94UG1  ALDKIEEEQA KSK.....K KAQQATADTR N.....SS.Q VSQNYPIVQN
F1_BE_VI85  ALDKLEEEQN KSQQ..... .KTQQAAADK .........G VSQNYPIVQN
F1_BR_93BR  ALEKLEEEQN KGRQ..... .KTQQATAEK .........G VSQNYPIVQN
F1_FI_FIN9  ALDKLEEEQN KSQQKT.... .QQAAAAADK .........G VSQNYPIVQN
F1_FR_MP41  ALEKLEEEQN KGQQ..... .KTQQAAADK .........G VSQNYPIVQN
F2_CM_MP25  ALDKLQEEQD KHQQ..... .KTQQATADK GVSK.....G VSQNYPILQN
F2KU_BE_VI  ALDKLEEEQK TQKQQ..... .KTQQEAAGK .........G VSQNYPIVQN
G_BE_DRCBL  ALEEVEKIQK KSQQ..... .......KEN S.....SS.Q VSQNYPIVQN
G_NG_92NG0  APEEVEKIQK NSQQE..... .IQQAAKNEG N.....SN.P VSQNYPIVQN
G_SE_SE616  ALEEVEKIQK KSQEK..... .IQQAAMDKG N.....SN.Q VSQNYPIVQN
H_BE_VI991  ALGKIEEIQN KNKQR..... .TQQAPAAAD K.EK...DSK ISQNYPIVQN
H_BE_VI997  ALGKIEEIQN KRQQK..... .TQQATANK. ..ER...DNK VSQNYPIVQN
H_CF_90CF0  ALDKIEEIQN KSQQK..... .TQQAAADK. ..EK...DNK VSQNYPIVQN
J_SE_SE702  ALDKIEEIQN KNKQQ..... .TQKAETDKK DNS.....Q VSQNYPIVQN
```

```
J_SE_SE788   ALEKIEEIQN KNKQQ..... .AQKAETDKK DNS......Q VSQNYPIVQN
K_CD_EQTB1   ALDKLEEEQ. NRTQQ..... .KTQQGKADK .........G VSQNYPIVQN
K_CM_MP535   ALDKLEEEQ. NKSQR..... .KTQQEAADK .........G VSQNYPIVQN
N_CM_YBF30   ALDKIKEKQE QHKPE....P KNPEAGAAAA TDS......N ISRNYPLVQT
O_CM_ANT70   AIQKLKEVMG SRKS...... .ADAAKEDTS A.....R..Q AGQNYPIVSN
O_CM_MVP51   AIQKLKEVMA SRKS...... .AEAAKEETS P.....R..Q TSQNYPIVTN
O_SN_99SE_   AIQKLKEVMG SRKS...... .AGTAKEDTS A.....R..Q TGQNYPIVTN
O_SN_99SE_   AIQKLKEVMG SRKS...... .AGAAKEDTS A.....R..Q TGQNYPVVAN
U_CD___83C   ALDKIEEVQN KGKQK..... .AQQAAADK. ..GS...NSQ VSQNYPIVQN

             151                                                  200
00BW0762_1   IQGQMVHQSL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW0768_2   LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW0874_2   LQGQMVHNPI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW1471_2   AQGQMVHQSI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW1616_2   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPSDLNTM
00BW1686_8   LQGQMVHQAI SPRTLNAWVK VIEEKGFNPE VIPMFTALSE GATPQDLNTM
00BW1759_3   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLHTM
00BW1773_2   LQGQMVHQPI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPSDLNTM
00BW1783_5   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW1795_6   LQGQMVHQAI SPRTLNAWVK VIEDKAFSPE VIPMFTALSE GATPGDLNTM
00BW1811_3   AQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW1859_5   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALPE GATPQDLNTM
00BW1880_2   LQGQMVHQPL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW1921_1   LQGQMVHQSI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW2036_1   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW2063_6   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE IIPMFTALSE GATPQDLNTM
00BW2087_2   AQGQMVHQTI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW2127_2   LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW2128_3   LQGQMVHQPL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW2276_7   IQGQMVHQPL SPRTLNAWVK VVEEKAFSPE VIPMFTALSE GATPQDSNTM
00BW3819_3   LQGQMVHQNI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW3842_8   LQGQMVHQAI SPRTLNAWVK VVEEKGFNPE VIPMFTALSE GATPQDLNTM
00BW3871_3   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW3876_9   LQGQMVHQSL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GAAPQDLNTM
00BW3886_8   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW3891_6   LQGQMVHQSL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
00BW3970_2   LQGQMVHQPI APRTLNAWVK VIEEKNFSPE IIPMFTALSE GATPQDLNTM
00BW5031_1   LQGQMVHQAI SPRTLNAWVK VIGEKAFNPE VIPMFTALSE GATPQDLNSM
 96BW01B21   LQGQMVHQPL SPRTLNAWVK VIKEKAFSPE IIPIFTALSE GATPRDLNTM
  96BW0407   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
  96BW0502   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
 96BW06_J4   LQGQMVHQPI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
 96BW11_06   LQGQMVHQAI SPRTLNAWVK VIEEKGFSPE VIPMFTALSE GATPQDLNTM
  96BW1210   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
 96BW15B03   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE IIPMFTALSE GATPQDLNTM
 96BW16_26   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
 96BW17A09   AQGQMVHQAI SPRTSNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
 96BWMO1_5   LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFTALSE GATPQDLNTM
 96BWMO3_2   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQILNTM
98BWMC12_2   VQGQMVHQPI SPRTLNAWVK VVQEKGFNPE VIPMFTALSE GATPQDLNTM
98BWMC13_4   LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLTTM
98BWMC14_a   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
98BWMO14_1   IQGQMVHQAI SPRTLNAWVK VIEEKGFNPE VIPMFTALSE GATPQDLNTM
98BWMO18_d   LQGQVVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
98BWMO36_a   LQGQMVHQAP SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
98BWMO37_d   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
99BW3932_1   LQGQMVHQSL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
99BW4642_4   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
99BW4745_8   LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
99BW4754_7   LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
```

```
99BWMC16_8  LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFTALSE GATPQDLNTM
A2_CD_97CD  AQGQMVHQAV SPRTLNAWVK VVEEKAFSPE VIPMFTALSE GATPQDLNTM
A2_CY_94CY  AQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFTALSE GATPQDLNTM
A2D___97KR  AQGQMTYQNL SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
A2G_CD_97C  AQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
A_BY_97BL0  AQGQMTHQSM SPRTLNAXVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
A_KE_Q23_A  AQGQMIHQSL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
A_SE_SE659  AQGQMIHQSL SPRTLNAWVK VIEEKGFNPE VIPMFSALSE GATPQDLNMM
A_SE_SE725  AQGQMVHQSL SPRTLNAWVK VIEEKAFSPE VIPVFSALSE GATPQDLNMM
A_SE_SE753  AQGQMVHQSL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
A_SE_SE853  AQGQMIHQNL SPRTLNAWVK VIEEKGFNPE VIPMFSALSE GATPQDLNMM
A_SE_SE889  AQGQMVHQSL SPRTLNAWVK VIEEKGFSPE VIPMFSALSE GATPQDLNTM
A_SE_UGSE8  AQGQMVHQVM SPRTLNAWVK VIEERAFSPE VIPMFSALSE GATPGDLNMM
A_UG_92UG0  AQGQMIHQSL SPRTLNAWVK VIEEKALSPE VIPMFSALSE GATPHDLNMM
A_UG_U455_  AQGQPVHQAL SPRTLNAWVK VVEDKAFSPE VIPMFSALSE GATPQDLNMM
AC_IN_2130  LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
AC_RW_92RW  AQGQMVHQAI SPRTLNAWVK VIEEKAFSPE IIPMFTALSE GATPQDLNTM
AC_SE_SE94  PQGQMVHQPI SPRTLNAWVK VIEDKAFSQE VIPMFTALSE GATPQDLNTM
ACD_SE_SE8  AQGQMVHQSL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
ACG_BE_VI1  AQEQMVHQSM SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
AD_SE_SE69  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
AD_SE_SE71  AQGQMIHQAL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GAAPQDLNTM
ADHK_NO_97  AQGQMVHQAI SPRTLNAWVK AIEEKAFSPE VIPMFSALSE GATPQDLNMM
ADK_CD_MAL  AQGQMIHQAI SPRTLNAWVK VIEEKAFSPE VIPMFSALSD GATPQDLNMM
AG_BE_VI11  AQGQMVHQSM SPRTLNAWVK VIEEKGFSPE VIPMFSALSE GATPQDLNMM
AG_NG_92NG  AQGQVVHQPI SPRTLNAWVK VIEEKNFSPE VIPMFSALSE GATPQDLNMM
AGHU_GA_VI  AQGQMVHQAM SPRTLNAWVK VIEEKNFSPE VIPMFTALSE GATPQDLNTM
AGU_CD_Z32  AQGQMVHQAL TPRTLNAWVK VVEEKGFSPE VIPMFSALSE GATPQDLNMM
AJ_BW_BW21  AQGQMVHQAV SPRTLNAWVK VVEEKNFSPE VIPMFSALSE GATPQDLNTM
B_AU_VH_AF  LQGQMVHQPI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
B_CN_RL42_  LQGQMVHQPI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_DE_D31_U  LQGQMVHQPI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_DE_HAN_U  LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_FR_HXB2_  IQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_GA_OYI___ LQGQMVHQPI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_GB_CAM1_  LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALAE GATPQDLNTM
B_GB_GB8_A  LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_GB_MANC_  LQGQMVHQSI SPRTLNAWVK VVEEKAFSPE VIPMFSALAE GATPQDLNTM
B_KR_WK_AF  LQGQMVHQPI SPRTLNARVK VIEEKAFSPE VIPMFSALSE GAIPQDLNTM
B_NL_3202A  LQGQMVHQAL SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_TW_TWCYS  LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_US_BC_L0  IQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_US_DH123  IQGQMVHQAL SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_US_JRCSF  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
B_US_MNCG_  IEGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_US_P896_  IQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_US_RF_M1  LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_US_SF2_K  LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
B_US_WEAU1  LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
B_US_WR27_  LQGQMVHQAL SPRTLFAWVK VVEEKAFSPE VIPMFSALSE GATPQDLYTM
B_US_YU2_M  LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
BF1_BR_93B  LQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
C_BR_92BR0  LQGQMVHQPI SARTLNAWVK VVEEKAFSPE VIPMFTALSE GATPQDLNTM
C_BW_96BW0  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
C_BW_96BW1  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
C_BW_96BW1  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE IIPMFTALSE GATPQDLNTM
C_BW_96BW1  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
C_ET_ETH22  MQGQMVHQPI SARTLNAWVK VVEEKAFSPE VIPMFTALSE GATPQDLNTM
C_IN_93IN1  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
C_IN_93IN9  LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
C_IN_93IN9  LQGQMVHQPI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
```

```
C_IN_94IN1  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
C_IN_95IN2  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
CRF01_AE_C  AQGQMVHQPL SPRTLNAWVK VVEEKGFNPE VIPMFSALSE GATPQDLNMM
CRF01_AE_C  AQGQMVHQAL SPRTLNAWVK VVEEKGFNPE VIPMFSALSE GATPQDLNMM
CRF01_AE_C  AQGQMIHQAL SPRTLNAWVK VVEEKGFNPE VIPMFSALSE GATPQDLNMM
CRF01_AE_T  AQGQMVHQPV SPRTLNAWVK VVEEKGFNPE VIPMFSALSE GATPQDLNMM
CRF01_AE_T  AQGQMVHQPL SPRTLNAWVK VVEEKGFNPE VIPMFSALSE GATPQDLNMM
CRF01_AE_T  AQGQMVHQPL SPRTLNAWVK VIEEKGFNPE VIPMFSALSE GATPQDLNMM
CRF01_AE_T  AQGQMVHQPI SPRTLNAWVK VIEEKGFSPE VIPMFSALSE GATPQDLNMM
CRF01_AE_T  AQGQMAHQPL SPRTLNAWVK VVEEKGFNPE VIPMFSALSE GATPQDLNMM
CRF01_AE_T  AQGQMVHQPV SPRTLNAWVK VVEEKGFNPE VIPMFSALSE GATPQDLNMM
CRF02_AG_F  AQGQMTHQPM SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF02_AG_F  AQGQMTHQPM SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNMM
CRF02_AG_G  AQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFTALSE GATPQDLNMM
CRF02_AG_N  AKGQMTHQSM SPRTLNAWVK VIEEKGFSPE VIPMFTALSE GATPQDLNMM
CRF02_AG_S  AQGQMTHQSI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF02_AG_S  AQGQMVHQSI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF03_AB_R  AQGQMTHQSM SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF03_AB_R  AQGQMTHQSM SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF04_cpx_  AQGQMVHQSI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF04_cpx_  AQGQMVHQAM SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF04_cpx_  AQGQMVHQSI SPRTLNAWVK VVEEKGFSPE VIPMFSALSE GATPQDLNMM
CRF05_DF_B  IQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF05_DF_B  LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
CRF06_cpx_  AQGQMVHQAM SPRTLNAWVK VIEDKAFSPE VIPMFTALSE GATPQDLNTM
CRF06_cpx_  AQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNMM
CRF06_cpx_  AQGQMIHQAI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF06_cpx_  AQGQMVHQAI SPRTLNAWVK AIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF11_cpx_  AQGQMVHQPV SPRTLNAWVK AIEEKAFSPE VIPMFSALSE GATPQDLNMM
CRF11_cpx_  AQGQMVHQAM SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNMM
D_CD_84ZR0  LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNMM
D_CD_ELI_K  LQGQMVHQAI SPRTLNAWVK VIEEKAFNPE VIPMFSALSE GATPQDLNTM
D_CD_NDK_M  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
D_UG_94UG1  LQGQMVHHPL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
F1_BE_VI85  LQGQMVHQSL SPRTLNAWVK VIEEKAFNPE VIPMFSALSE GATPQDLNTM
F1_BR_93BR  LQGQMVHQSL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPTDLNTM
F1_FI_FIN9  LQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
F1_FR_MP41  LQGQMVHQPI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
F2_CM_MP25  LQGQMVHQSL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
F2KU_BE_VI  LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
G_BE_DRCBL  AQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFTALSE GATPQDLNTM
G_NG_92NG0  AQGQMIHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNTM
G_SE_SE616  AQGQMVHQAI TPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNLM
H_BE_VI991  AQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNAM
H_BE_VI997  AQGQMVHQPI SXRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNAM
H_CF_90CF0  AQGQMVHQAI SPRTLNAWVK VVEEKAFSPE VIPMFSALSE GATPQDLNAM
J_SE_SE702  LQGQPVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
J_SE_SE788  LQGQPVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
K_CD_EQTB1  LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM
K_CM_MP535  LQGQMVHQAL SPRTLNAWVK VIEEKAFSPE VIPMFTALSE GATPQDLNTM
N_CM_YBF30  AQGQMVHQPL TPRTLNAWVK VIEEKAFSPE VIPMFMALSE GATPSDLNTM
O_CM_ANT70  AQGQMVHQAI SPRTLNAWVK AVEEKAFNPE IIPMFMALSE GAISYDINTM
O_CM_MVP51  AQGQMVHQAI SPRTLNAWVK AVEEKAFNPE IIPMFMALSE GAVPYDINTM
O_SN_99SE_  AQGQMVHQSL SPRTLNAWVK AVEEKAFNPE IIPMFMALSE GAIPYDTNTM
O_SN_99SE_  AQGQMVHQSL SPRTLNAWVK AVEEKAFNPE IIPMFMALSE GAIPYDTNTM
U_CD___83C  MQGQMVHQAI SPRTLNAWVK VIEEKAFSPE VIPMFSALSE GATPQDLNTM

            201                                                 250
00BW0762_1  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
00BW0768_2  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PAHAGPVAPG QMREPRGSDI
00BW0874_2  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPVAPG QMREPRGSDI
```

```
00BW1471_2   LNTVGGHQ.A AMQMLKDTIN EEVAEWDRLH PVQAGPIAPG QMRDPRGSDI
00BW1616_2   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRMH PVQAGPVAPG QMRDPRGSDI
00BW1686_8   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
00BW1759_3   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
00BW1773_2   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
00BW1783_5   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVQAGPVAPG QIREPRGSDI
00BW1795_6   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
00BW1811_3   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
00BW1859_5   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIAPG QMREPRGSDI
00BW1880_2   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMRDPRGSDI
00BW1921_1   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QLREPRGSDI
00BW2036_1   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPNPAG QMREPRGSDI
00BW2063_6   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPAAPG QMREPRGSDI
00BW2087_2   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
00BW2127_2   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
00BW2128_3   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIAPG QMREPRGSDI
00BW2276_7   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
00BW3819_3   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
00BW3842_8   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVQAGPVAPG QIREPRGSDI
00BW3871_3   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
00BW3876_9   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRIH PVHAGPVAPG QMRDPRGSDI
00BW3886_8   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
00BW3891_6   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
00BW3970_2   LNAVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QLREPRGSDI
00BW5031_1   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
 96BW01B21   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRTH PVHAGPVAPG QLREPRGSDI
  96BW0407   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIAPG QMREPRGSDI
  96BW0502   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVQAGPVAPG QMRDPRGSDI
 96BW06_J4   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVQAGPVAPG QMRDPRGSDI
 96BW11_06   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
  96BW1210   LNTVGGHQ.A AMQMLKDTIN EEAAGWDRLH PVHAGPVAPG QMREPRGSDI
 96BW15B03   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
 96BW16_26   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIAPG QMREPRGSDI
 96BW17A09   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
 96BWMO1_5   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
 96BWMO3_2   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRMH PVHAGPVAPG QMREPRGSDI
98BWMC12_2   LNTVGGRQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
98BWMC13_4   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVQAGPVAPG QIREPRGSDI
98BWMC14_a   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMRDPRGSDI
98BWMO14_1   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRTH PVHAGPVAPG QMREPRGSDI
98BWMO18_d   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
98BWMO36_a   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
98BWMO37_d   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVQAGPIAPG QIREPRGSDI
99BW3932_1   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRTH PVHAGPVAPG QMREPRGSDI
99BW4642_4   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
99BW4745_8   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRIH PVHAGPVAPG QMREPRGSDI
99BW4754_7   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIAPG QIREPRGSDI
99BWMC16_8   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRTH PVHAGPIAPG QMREPRGSDI
A2_CD_97CD   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVQAGPIPPG QMREPRGSDI
A2_CY_94CY   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
A2D___97KR   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
A2G_CD_97C   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGS..
A_BY_97BL0   LNIVGGHQ.A AMQMLKDTIN EEAAXXDRLH PAQAGPFPPG QMREPRGSDI
A_KE_Q23_A   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
A_SE_SE659   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGS..
A_SE_SE725   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PAHAGPVAPG QMREPRGS..
A_SE_SE753   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
A_SE_SE853   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGS..
A_SE_SE889   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGS..
A_SE_UGSE8   LNIVGGHQ.A AMEMLKDTIN EEAAEWDRTH PIHAGPVAPG QMREPRGSDI
A_UG_92UG0   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
```

```
A_UG_U455_   LNVVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
AC_IN_2130   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PAQAGPIAPG QMREPRGSDI
AC_RW_92RW   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVQAGPVAPG QIREPRGSDI
AC_SE_SE94   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PAQAGPVAPG QMREPRGS..
ACD_SE_SE8   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRIH PVHAGLIAPG QMREPRGSDI
ACG_BE_VI1   LNIVGGHQ.A AMQMLKDTIN EEAADWDRTH PVHAGPNPPG QMREPRGSDI
AD_SE_SE69   LSTVGGHQ.A AMQILKETIN EEAADWDRLH PVHAGPNAPG QMREPRGSDI
AD_SE_SE71   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGS..
ADHK_NO_97   LNIVGGHQ.A AMQMLKDTIN EEAADWDRLH PVHAGPIPPG QMREPRGSDI
ADK_CD_MAL   LNIVGGHQ.A AMQMLKDTIN EEAADWDRVH PVHAGPIPPG QMREPRGSDI
AG_BE_VI11   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
AG_NG_92NG   LNTVGGHQ.A AMQMLKDSIN EEAAEWDRLH PQQAGPIPPG QIREPRGSDI
AGHU_GA_VI   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRIH PVQAGPIPPG QIREPRGSDI
AGU_CD_Z32   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PPQAGPIPPG QIREPRGSDI
AJ_BW_BW21   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
B_AU_VH_AF   LNTVGGHQAA AMQMLKETIN EEAADWDRLH PVHAGPIAPG QMREPRGSDI
B_CN_RL42_   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
B_DE_D31_U   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_DE_HAN_U   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_FR_HXB2_   LNTVGGHQ.A AMQMLKETIN EEAAEWDRVH PVHAGPIAPG QMREPRGSDI
B_GA_OYI__   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_GB_CAM1_   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_GB_GB8_A   LNTVGGHQ.A AMQMLKETIN EEAAEWDRVH PVHAGPVAPG QMREPRGSDI
B_GB_MANC_   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
B_KR_WK_AF   LNTVGGHQ.A AMQMLKETIN EEAADWDRLH PVHAGPIAPG QMREPRGSDI
B_NL_3202A   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_TW_TWCYS   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_US_BC_L0   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVQAGPVAPG QMREPRGSDI
B_US_DH123   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_US_JRCSF   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_US_MNCG_   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPITPG QMREPRGSDI
B_US_P896_   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVQAGPVAPG QMREPRGSDI
B_US_RF_M1   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_US_SF2_K   LNTVGGHQ.A AMQMLKETIN EEAAEWDRVH PVHAGPIAPG QMREPRGSDI
B_US_WEAU1   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
B_US_WR27_   LNTVGGHQ.A AMQMLKETIN DEAAEWDRLH PVQAGPVAPG QMREPRGSDI
B_US_YU2_M   LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
BF1_BR_93B   LNTVGGHQ.A AMQMLKETIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
C_BR_92BR0   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
C_BW_96BW0   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
C_BW_96BW1   LTTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMREPRGSDI
C_BW_96BW1   LNTVGGHQ.A AMQMLKDTIN EEAAGWDRLH PVHAGPVAPG QMREPRGSDI
C_BW_96BW1   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
C_ET_ETH22   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QMRDPRGSDI
C_IN_93IN1   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PIHAGPIAPG QMREPRGSDI
C_IN_93IN9   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PIHAGPIAPG QMREPRGSDI
C_IN_93IN9   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QIREPRGSDI
C_IN_94IN1   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRIH PVHAGPIAPG QMREPRGSDI
C_IN_95IN2   LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVPAGPIAPG QLREPRGSDI
CRF01_AE_C   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
CRF01_AE_C   LNIVGGHQ.A AMQILKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGADI
CRF01_AE_C   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF01_AE_T   LNIVGEHQ.A AMQMLKETIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF01_AE_T   LNIVGGHQ.A AMQMLKETTN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF01_AE_T   LNIVGGHQ.A AMQMLKETIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF01_AE_T   LNIVGGHQ.A AMQMLKETIN EEPAEWDRVH PVHAGPIPPG QIREPRGSDI
CRF01_AE_T   LNIVGGHQ.A AMQMLKETIN EEPAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF01_AE_T   LNIVGGHQ.A AMQMLKETIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF02_AG_F   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF02_AG_F   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF02_AG_G   LNIVGGHQ.A AMQMLKDTIN EEAAEWDRTH PIHAGPNPPG QMREPRGSDI
```

```
CRF02_AG_N  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF02_AG_S  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF02_AG_S  LNIVGGHQ.A AMQMLKDTIN EEAADWDRTH PVHAGPIPPG QMREPRGSDI
CRF03_AB_R  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PAQAGPFPPG QMREPRGSDI
CRF03_AB_R  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PAQAGPFPPG QMREPRGSDI
CRF04_cpx_  LNIVGGHQ.A AMQMLKDTIN EEAADWDRTH PVHAGPIPPG QMREPRGSDI
CRF04_cpx_  LNIVGGHQ.A AMQMLKDTIN EEASEWDRAH PVHAGPIPPG QMREPRGSDI
CRF04_cpx_  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PAHAGPNPAG QMREPRGSDI
CRF05_DF_B  LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVQAGPVAPG QMRDPRGSDI
CRF05_DF_B  LNTVGGHQ.A AMQMLKETIN EEAAEWDRVH PAQAGPIAPG QIREPRGSDI
CRF06_cpx_  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QIREPRGSDI
CRF06_cpx_  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
CRF06_cpx_  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRMH PVQAGPMPPG QMREPRGSDI
CRF06_cpx_  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PAQAGPIPPG QIRDPRGSDI
CRF11_cpx_  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVQAGPIAPG QMRDPRGSDI
CRF11_cpx_  LNIVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPAPPG QMREPRGSDI
D_CD_84ZR0  LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVQAGPVAPG QMREPRGSDI
D_CD_ELI_K  LNTVGGHQ.A AMQMLKETIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
D_CD_NDK_M  LNTVGGHQ.A AMQMLKETIN DEAAEWDRLH PVHAGPVAPG QMREPRGSDI
D_UG_94UG1  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPVAPG QLREPRGSDI
F1_BE_VI85  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPAPPG QMREPRGSDI
F1_BR_93BR  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PTQAGPIPPG QIREPRGSDI
F1_FI_FIN9  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
F1_FR_MP41  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PAHAGPILPG QMREPRGSDI
F2_CM_MP25  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
F2KU_BE_VI  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIAPG QMREPRGSDI
G_BE_DRCBL  LNTVGGHQ.A AMQMLKETIN DEAAEWDRLH PQQAGPIAPG QIRDPTGSDI
G_NG_92NG0  LNTVGGHQ.A AMQMLKDTIN DEAAEWDRIH PQQAGPIPPG QIREPSGSDI
G_SE_SE616  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRMH PQQAGPFPPG QIREPRGSDI
H_BE_VI991  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
H_BE_VI997  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRLH PVHAGPIPPG QMREPRGSDI
H_CF_90CF0  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIPPG QMREPRGSDI
J_SE_SE702  LNTIGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPVAPG QVREPRGSDI
J_SE_SE788  LNTIGGHQ.A AMQMLKDTIN EEAAEWDRVH PVHAGPIAPG QVREPRGSDI
K_CD_EQTB1  LNTVGGHQ.A AMQMLKDTIN EEAAEWDRMH PVQAGPIPPG QIREPRGSDI
K_CM_MP535  LNTVGGHQ.A AMQMLKDTIN DEAAEWDRLH PVHAGPIPPG QMREPRGSDI
N_CM_YBF30  LNTVGGHQ.A AMQMLKEVIN EEAADWDRTH PVPVGPLPPG QLRDPRGSDI
O_CM_ANT70  LNAIGGHQ.G ALQVLKEVIN EEAVEWDRTH PPPVGPLPPG QIREPTGSDI
O_CM_MVP51  LNAIGGHQ.G ALQVLKEVIN EEAAEWDRTH PPAMGPLPPG QIREPTGSDI
O_SN_99SE_  LNAIGGHQ.G ALQVLKEVIN EEAAEWDRTH PPAAGPLPVG QIREPTGSDI
O_SN_99SE_  LNAIGGHQ.G ALQVLKEVIN EEAAEWDRTH PQAAGPLPPG QIREPTGSDI
U_CD___83C  LNTVGGHQ.A AMQMLKDTIN EEAADWDRLH PVHAGPIPPG QMREPRGSDI
```

```
          251                                                        300
00BW0762_1  AGTTSTLQEQ IAWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
00BW0768_2  AGTTSNLQEQ IAWMTA.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
00BW0874_2  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW1471_2  AGTTSTLQEQ VAWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW1616_2  AGTTSTLQEQ IAWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
00BW1686_8  AGTTSNLQEQ VAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW1759_3  AGTTSTLQEQ IAWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW1773_2  AGTTSTLQEQ ITWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW1783_5  AGTTSTLQEQ ITWMTS.NPP IPVGDIYKRW IVLGLNKIVR MYSPVSILDI
00BW1795_6  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW1811_3  AGTTSTLQEQ IAWMTN.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW1859_5  AGTTSTLQEQ IAWMTG.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW1880_2  AGTTSTLQEQ ITWMTS.NPP IPVGDIYKRW IVLGLNKIVR MYSPVSILDI
00BW1921_1  AGTTSTLQEQ IAWMTN.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW2036_1  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW2063_6  AGTTSTLAEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW2087_2  AGTTSTLQEQ IAWMTN.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW2127_2  AGTTTTLQEQ INWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
00BW2128_3  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW2276_7  AGTTSTLQEQ IAWMTS.NPA IPVGDIYKRW IILGLDKIVR MYSPVSILDI
00BW3819_3  AGTTSTLQEQ VAWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW3842_8  AGTTSNLQEQ ITWMTS.NPP IPVGEIYKRW IVLGLNKIVR MYSPVSILDI
00BW3871_3  AGTTSTLQEQ IDWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
00BW3876_9  AGTTSTLQEQ IAWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
00BW3886_8  AGTTSTLQEQ IAWMTN.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW3891_6  AGTTSNLQEQ INWMTA.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
00BW3970_2  AGTTSTLQEQ IAWMTN.NPP VPVGDIYKRW IVLGLNKIVR MYSPVSILDI
00BW5031_1  AGTTSTLQEQ IAWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
 96BW01B21  AGTTSNLQEQ IAWMTA.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
  96BW0407  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
  96BW0502  AGATSTLQEQ IAWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
 96BW06_J4  AGTTSTLQEQ IGWMTH.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
 96BW11_06  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
  96BW1210  AGTTSNLQEQ INWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
 96BW15B03  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
 96BW16_26  AGTTSTLQEQ IAWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
 96BW17A09  AGTTSTLQEQ IAWMTN.NPP IPVGDIYKRW ITMGLNKIVR MYSPVSILDI
 96BWMO1_5  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
 96BWMO3_2  AGSTSTLQEQ IAWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
98BWMC12_2  AGSTSNLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
98BWMC13_4  AGTTSTLQEQ IAWMTR.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
98BWMC14_a  AGTTSTLQEQ VGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
98BWMO14_1  AGTTSTLQEQ IAWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
98BWMO18_d  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
98BWMO36_a  AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IIMGLNKIVR MYSPVSILDI
98BWMO37_d  AGTNSTLQEQ IAWMTN.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
99BW3932_1  AGTTSTLQEQ IQWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
99BW4642_4  AGTTSTLQEQ VTWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
99BW4745_8  AGTTSTLQEQ INWMTG.NPP IPVGDIYKRW IIMGLNKIVR MYSPVSILDI
99BW4754_7  AGTTSTLQEQ ITWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
99BWMC16_8  AGTTSTLQEQ IAWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
A2_CD_97CD  AGATSNLQEQ IGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
A2_CY_94CY  AGTTSTLQEQ IGWMTS.DPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
A2D___97KR  AGTTSTLQEQ IGWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
A2G_CD_97C  .......... .......... .......... .......... ..........
A_BY_97BL0  AGTTSTLQEQ IGXMTS.NPP IPXGDIYKRX IILGLNKIVR MYSPVSILDI
A_KE_Q23_A  AGTTSTPQEQ IGWMTG.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
A_SE_SE659  .......... .......... .......... .......... ..........
A_SE_SE725  .......... .......... .......... .......... ..........
A_SE_SE753  AGTTSTPQEQ IGWMTG.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
```

```
A_SE_SE853   .......... .......... .......... .......... ..........
A_SE_SE889   .......... .......... .......... .......... ..........
A_SE_UGSE8   AGTTSTLQEQ IAWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
A_UG_92UG0   AGTTSTPQEQ IAWMTG.NPP IPVGDIYKRW MILGLNKIVR MYSPVSILDI
A_UG_U455_   AGTTSTVQEQ IGWMTG.NPP IPVGDIYRRW IILGLNKIVR MYSPVSILDI
AC_IN_2130   AGTTSTLQEQ IAWMTG.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
AC_RW_92RW   AGTTSTLQEQ IAWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
AC_SE_SE94   .......... .......... .......... .......... ..........
ACD_SE_SE8   AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
ACG_BE_VI1   AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
AD_SE_SE69   AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
AD_SE_SE71   .......... .......... .......... .......... ..........
ADHK_NO_97   AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
ADK_CD_MAL   AGTTSTLQEQ IGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
AG_BE_VI11   AGSTSTLQEQ VGWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
AG_NG_92NG   AGTTSTLQEQ ITWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
AGHU_GA_VI   AGTTSTTQEQ IGWMTG.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
AGU_CD_Z32   AGTTSTLQEQ IRWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
AJ_BW_BW21   AGTTSTLQEQ IGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
B_AU_VH_AF   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_CN_RL42_   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTCILDI
B_DE_D31_U   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_DE_HAN_U   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_FR_HXB2_   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_GA_OYI__   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_GB_CAM1_   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_GB_GB8_A   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
B_GB_MANC_   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_KR_WK_AF   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPASILDI
B_NL_3202A   AGTTSTLQEQ IGWMTH.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_TW_TWCYS   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_US_BC_L0   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPSSILDI
B_US_DH123   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IIMGLNKIVR MYSPTSILDI
B_US_JRCSF   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
B_US_MNCG_   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPSSILDI
B_US_P896_   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPSSILDI
B_US_RF_M1   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPISILDI
B_US_SF2_K   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_US_WEAU1   AGTTSTLQEQ IAWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
B_US_WR27_   AGXTSTLXXX IGWMTH.XPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
B_US_YU2_M   AGTTSTLQEQ IGWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPTSILDI
BF1_BR_93B   AGTTSTLQEQ IQWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPTSILGI
C_BR_92BR0   AGTTSTLQEQ ITWMTN.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
C_BW_96BW0   AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
C_BW_96BW1   AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
C_BW_96BW1   AGTTSNLQEQ INWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
C_BW_96BW1   AGTTSTLQEQ IAWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
C_ET_ETH22   AGTTSTLQEQ IAWMTG.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
C_IN_93IN1   AGTTSSLQEQ IAWMTG.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
C_IN_93IN9   AGTTSSLQEQ IAWMTG.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
C_IN_93IN9   AGTTSTLQEQ IAWMTG.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
C_IN_94IN1   AGTTSTLQEQ IAWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDT
C_IN_95IN2   AGTTSTLQEQ IAWMTN.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF01_AE_C   XGTTSNLQEQ IGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF01_AE_C   AGTTSTLHEQ IGWMTS.NPP IPVGEIYKKW IILGLNKIVR MYSPVSILDI
CRF01_AE_C   AGTTSTLQEQ IGWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
CRF01_AE_T   AGTTSTLQEQ IGWMTN.NPP IPVGDIYKRW IILGLNKIVR MYRPVSILDI
CRF01_AE_T   AGTTSTLQEQ IGWMTN.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF01_AE_T   AGTTSTLQEQ IGWMTN.NPP IPVGDIYKRW IILGLNKIVR MYQPVSILDI
CRF01_AE_T   AGTTTTLQEQ IGWMTN.NPP IPVGSIYKRW IILGLNKIVR MYSPVSILDI
CRF01_AE_T   AGTTSTLQEQ IGWMTN.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
```

```
CRF01_AE_T  AGTTSTLQEQ IGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF02_AG_F  AGTTSTPQEQ IGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF02_AG_F  AGTTSTLQEQ IGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF02_AG_G  AGTTSNLQEQ IAWMTG.NPP IPVGEIYKRW IVLGLNKIVR MYSPVGILDI
CRF02_AG_N  AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IVLGLNKIVR MYSPVGILDI
CRF02_AG_S  AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IVLGLNKIVR MYSPVSILDI
CRF02_AG_S  AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IVLGLNKIVR MYSPVSILDI
CRF03_AB_R  AGSSSTLQEQ IGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF03_AB_R  AGTTSTLQEQ IGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF04_cpx_  AGTTSTLQEQ IGWMTS.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF04_cpx_  AGTTSTLQEQ IGWMTS.NPP VPVGEIYKRW IILGLNKIVR TYSPISILDI
CRF04_cpx_  AGTTSTLQEQ VGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
CRF05_DF_B  AGTTSTLQEQ IAWMTN.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
CRF05_DF_B  AGTTSTLQEQ ITWMTN.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
CRF06_cpx_  AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
CRF06_cpx_  AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVGILDI
CRF06_cpx_  AGTTSTLLEQ IGWMTG.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
CRF06_cpx_  AGTTSTLQEQ IRWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
CRF11_cpx_  AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
CRF11_cpx_  AGTTGNLQEQ IGWMTG.NPP IPVGDIYRRW IILGLNKIVR MYSPVSILDI
D_CD_84ZR0  AGTTSTLQEQ IGWMTG.NPP IPVGEIYRRW IILGLNKIVR MYSPVSILDI
D_CD_ELI_K  AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
D_CD_NDK_M  AGTTSTLQEQ IAWMTS.NPP IPVGEIYKRW IIVGLNKIVR MYSPVSILDI
D_UG_94UG1  AGTTSNLQEQ IAWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
F1_BE_VI85  AGTTSTLQEQ IGWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
F1_BR_93BR  AGTTSTLQEQ IQWMTG.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
F1_FI_FIN9  AGTTSTLQEQ IQWMTG.NPP VPVGEMYKRW IILGLNKIVR MYSPVGILDI
F1_FR_MP41  AGTTSTLQEQ IQWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
F2_CM_MP25  AGTTSTLQEQ IQWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
F2KU_BE_VI  AGATSNLQEQ IAWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
G_BE_DRCBL  AGATSTLQEQ IAWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
G_NG_92NG0  AGTTSTLQEQ IRWMTS.NPP VPVGEIYKRW IILGLNKIVR MYSPVSILDI
G_SE_SE616  AGTTSTLQEQ IRWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
H_BE_VI991  AGTTSSLQEQ ITWMTG.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
H_BE_VI997  AGTTSTLQEQ VAWMTG.NPP IPVGDIYKRW IILGLNKIVR MYSPVSILDI
H_CF_90CF0  AGTTSTLQEQ IAWMTG.NPS IPVGDIYKRW IILGLNKIVR MYSPVSILDI
J_SE_SE702  AGTTSTLQEQ IAWMTG.NPA IPVGDIYKRW IILGLNKIVR MYSPVSILDI
J_SE_SE788  AGTTSNLQEQ IGWMTG.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
K_CD_EQTB1  AGTTSTLQEQ IGWMTG.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
K_CM_MP535  AGTTSTLQEQ ITWMTS.NPP IPVGEIYKRW IILGLNKIVR MYSPVSILDI
N_CM_YBF30  AGTTSTLAEQ IAWMTS.NPP VPVGDIYKRW IILGLNKIVR MYSPVSILDI
O_CM_ANT70  AGTTSTQQEQ VAWMTA.NPP VPVGDIYRRW IVLGLNRIVR MYSPVSILEI
O_CM_MVP51  AGTTSTQQEQ IHWTTRPNQP IPVGDIYRKW IVLGLNKMVK MYSPVSILDI
O_SN_99SE_  AGTTSTQQEQ IIWTTRGANS IPVGDIYRKW IVLGLNKMVK MYSPVSILDI
O_SN_99SE_  AGTTSTQQEQ VHWITRPNQP IPVGDIYRKW IVLGLNKVVK MYSPVSILDI
U_CD___83C  AGTTSTLQEQ VHWITRPNQP IPVGDIYRKW IVLGLNKVVK MYSPVSILDI
                                  301                          350
00BW0762_1  KQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILR
00BW0768_2  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW0874_2  KQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILR
00BW1471_2  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW1616_2  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW1686_8  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKAILR
00BW1759_3  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLIQN ANPDCKTILK
00BW1773_2  KQGPKEPFRD YVDRFFKVLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW1783_5  KQGPKEPFRD YVDRFFKTLR AEQSTQEVKN WMTDTLLIQN ANPDCKTILR
00BW1795_6  KQGPKEPFRD YVDRFFKTLR AEQSTQEVKN WMTDTLLVQN ANPDCKTILK
00BW1811_3  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKTILR
00BW1859_5  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW1880_2  KQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILR
```

```
00BW1921_1  KQGPKEPFRD YVDRFFKTLR AEQSSQEVKN WMTDTLLVQN ANPDCKTILR
00BW2036_1  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW2063_6  KQGPKEPFRD YVDRFFKTLR AEQSTQEVKN WVTDTLLVQN ANPDCKTILR
00BW2087_2  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILK
00BW2127_2  KQGPKEPFRD YVDRIFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILK
00BW2128_3  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKTILR
00BW2276_7  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTKTLLVQN ANPDCKTILR
00BW3819_3  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW3842_8  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW3871_3  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKTILR
00BW3876_9  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTGTLLVQN ANPDCKTILR
00BW3886_8  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILK
00BW3891_6  KQGPKEPFRD YVDRFFRTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW3970_2  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
00BW5031_1  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
 96BW01B21  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN GNPDCKNILR
  96BW0407  KQGPKEPFRD YVDRFFKTLR AEQSTQEVKN WMTDTLLVQN ANPDCKTILR
  96BW0502  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
 96BW06_J4  KQGPKESFRD YVDRFFKTLR AEQCTQDVKN WMTETLLVQN ANPDCKTILR
 96BW11_06  KQGPKESFRD YVDRFFKTLR AEQSSQEVKN WMTDTLLVQN ANPDCKTILR
  96BW1210  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCRTILK
 96BW15B03  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
 96BW16_26  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKPILR
 96BW17A09  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
 96BWMO1_5  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
 96BWMO3_2  KQGPKEPFRD YVDRFFKTLR AEQSSQEVKN WMTDTLLVQN ANPDCKTILR
98BWMC12_2  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
98BWMC13_4  KQGPKESFRD YVDRFFKTLR AEQATQEVKG WMTDTLLIQN ANPDCKTILR
98BWMC14_a  KQGPKEPFRD YVDRFFKTLR AEQATREVKN WMTDTLLVQN ANPDCKTILR
98BWMO14_1  KQGPKEPFRD YVDRFFKCLR AEQATQEVKD WMTETLLVQN ANPDCKTILR
98BWMO18_d  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
98BWMO36_a  KQGPKEPFRD YVDRFFKTLR AEQSTQEVKN WMTDTLLVQN ANPDCKTILR
98BWMO37_d  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILR
99BW3932_1  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILR
99BW4642_4  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKTILR
99BW4745_8  KQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILR
99BW4754_7  KQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMADTLLVQN ANPDCKTILR
99BWMC16_8  KQGPKEPFRD YVDRFFKTLR AEQSTQDVKN WMTDTLLVQN ANPDCKTILR
A2_CD_97CD  KQGPKEPFRD YVDRFFKTLR AEQASQDVKN WMTDTLLVQN ANPDCKTILR
A2_CY_94CY  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKSILR
A2D___97KR  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCRSILR
A2G_CD_97C  .......... .......... .......... .......... ANPDCKSILR
A_BY_97BL0  RQGPKEPFRD YVDRFFKTLX AEQATQDVKN WMTETLLVQN ANPDCKTILR
A_KE_Q23_A  KQGPKEPFRD YVDRFFKTFR AEQATQDVKN WMTDTLLVQN ANPDCKAILR
A_SE_SE659  .......... .......... .......... .......... ..........
A_SE_SE725  .......... .......... .......... .......... ..........
A_SE_SE753  KQGPKEPFRD YVDRFFKTLR AEQASQDVKN WMTETLLIQN ANPDCKSILR
A_SE_SE853  .......... .......... .......... .......... ..........
A_SE_SE889  .......... .......... .......... .......... ..........
A_SE_UGSE8  KQGPKEPFRD YVDRFFKALR AEQATQEVKG WMTDTLLVQN ANPDCKSILR
A_UG_92UG0  KQGPKEPFRD YVDRFFKTLR AEQATQEVKG WMTETLLIQN ANPDCKSILR
A_UG_U455_  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKSILR
AC_IN_2130  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
AC_RW_92RW  KQGPKEPFRD YVDRFFKTLR AEQASQDVKN WMTDTLLVQN ANPDCKTILR
AC_SE_SE94  .......... .......... .......... .......... ..........
ACD_SE_SE8  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKTILK
ACG_BE_VI1  RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCRTILR
AD_SE_SE69  RQGPKEPFRD YVDRFYKTLR AEQASRDVKN WMTETLLVQN ANPDCKTILK
AD_SE_SE71  .......... .......... .......... .......... ..........
ADHK_NO_97  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILK
ADK_CD_MAL  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKTILK
```

```
AG_BE_VI11   RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
AG_NG_92NG   KQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILR
AGHU_GA_VI   RQGPKEPFRE YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKTILR
AGU_CD_Z32   KQGPKEPFRD YVDRFFKTLR AEQATQEVKG WMTDTLLVQN ANPDCKTILR
AJ_BW_BW21   RQGPKEPFRD YVDRFFKTLR AEQASQDVKN WMTETLLVQN ANPDCKTILK
B_AU_VH_AF   KQGPKEPFRD YVDRFYRTLR AEQASQEVKN WMTETLLVQN ANPDCRTILK
B_CN_RL42_   RQGPKEPFRD YVDRFYKTLR AEQASQDVKN WMTETLLVQN ANPDCKTILK
B_DE_D31_U   RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKTILK
B_DE_HAN_U   KQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKTILK
B_FR_HXB2_   RQGPKEPFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
B_GA_OYI__   RQGPKEPFRD YVDRFYKTLR AEQASQDVKN WMTETLLVQN ANPDCKTILK
B_GB_CAM1_   RQGPKEPFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
B_GB_GB8_A   RQGPKEPFRD YVDRFYKTLR AEQASQDVKN WMTETLLVQN ANPDCKTILK
B_GB_MANC_   RQGPKEPFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
B_KR_WK_AF   RQGPKEPFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
B_NL_3202A   RQGPKESFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
B_TW_TWCYS   RQGPKEPFRD YVDRFYKTLR TEQASQDVKN WMTETLLVQN ANPDCKTILK
B_US_BC_L0   RQGPKESFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
B_US_DH123   RQGPKEPFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN SNPDCKTILK
B_US_JRCSF   RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKTILK
B_US_MNCG_   RQGPKEPFRD YVDRFYKTLR AEQASQEVKN RTTETLLVQN ANPDCKTILK
B_US_P896_   KQGPKEPFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
B_US_RF_M1   RQGPKEPFRD YVDRFYKTLR AEQASQDVKN WMTETFLVQN ANPDCKTILK
B_US_SF2_K   RQGPKEPFRD YVDRFYKTLR AEQASQDVKN WMTETLLVQN ANPDCKTILK
B_US_WEAU1   RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKTILK
B_US_WR27_   RQGPKEPFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
B_US_YU2_M   RQGPKEPFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
BF1_BR_93B   RQGPKEPFRD YVDRFYKTLR AEQTSQDVKN WMTETLLVQN ANPDCKTILK
C_BR_92BR0   KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
C_BW_96BW0   KQGPKEPFRD YVDRFFKTLR AEQSTQEVKN WMTDTLLVQN ANPDCKTILR
C_BW_96BW1   KQGPKESFRD YVDRFFKTLR AEQSSQEVKN WMTDTLLVQN ANPDCRTILK
C_BW_96BW1   KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
C_BW_96BW1   RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKPILR
C_ET_ETH22   KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
C_IN_93IN1   RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
C_IN_93IN9   KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
C_IN_93IN9   KQGPKEPFRD YVDRFFRTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
C_IN_94IN1   KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTDTLLVQN ANPDCKTILR
C_IN_95IN2   RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKTILR
CRF01_AE_C   RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF01_AE_C   RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF01_AE_C   RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF01_AE_T   RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF01_AE_T   RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF01_AE_T   RQGPKEPFRD YVDRFYKVLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF01_AE_T   RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF01_AE_T   RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF01_AE_T   RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF02_AG_F   RQGPKEPFRD YVDRFFKTLR AEQATQEVKN RMTETLLVQN ANPDCKSILR
CRF02_AG_F   RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILR
CRF02_AG_G   RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKTILR
CRF02_AG_N   RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN ANPDCKSILR
CRF02_AG_S   RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKSILR
CRF02_AG_S   RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTETLLVQN SNPDCRTILK
CRF03_AB_R   RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKTILR
CRF03_AB_R   RQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKTILR
CRF04_cpx_   RQGPKEPFRD YVDRFFKCLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF04_cpx_   RQGPKEPFRD YVDRFFKCLR AEQATQEVKN WMTETLLVQN ANPDCKSILK
CRF04_cpx_   RQGTKEPFRD YVDRFFKCLR AEQATQDVKN WMTETLLVQN ANPDCKSILK
CRF05_DF_B   RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKTILK
CRF05_DF_B   RQGPKEPFRD YVDRFYKTLR AEQASQDVKN WMTETLLVQN ANPDCKTILK
```

```
CRF06_cpx_  KQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILR
CRF06_cpx_  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILK
CRF06_cpx_  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILK
CRF06_cpx_  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILK
CRF11_cpx_  KQGPKEPFRD YVDRFFKALR AEQATQEVKG WMTDTLLIQN ANPDCKSILR
CRF11_cpx_  RQGPKEPFRD YVDRFFKTLR AEQASQEVKN WMTETLLVQN ANPDCKSILR
D_CD_84ZR0  RQGPKEPFRD YVDRFYKTLR AEQASQEVKN WMTETLLVQN ANPDCKTILK
D_CD_ELI_K  RQGPKEPFRD YVDRFYKTLR AEQASQDVKN WMTETLLVQN ANPDCKTILK
D_CD_NDK_M  RQGPKEPFRD YVDRFYKTLR AEQASQDVKN WMTETLLVQN ANPDCKTILK
D_UG_94UG1  RQGPKEPFRD YVDRFYKTLR AEQASQDVKN WMTETLLVQN ANPDCKTILK
F1_BE_VI85  KQGPKEPFRD YVDRFFKVLR AEQASQDVKG WMTDTLLVQN ANPDCKTILK
F1_BR_93BR  RQGPKEPFRD YVDRFFKTLR AEQATQEVKG WMTDTLLVQN ANPDCKTILK
F1_FI_FIN9  RQGPKEPFRD YVDRFFKALR AEQATQEVKG WMTDTLLVQN ANPDCKTILK
F1_FR_MP41  RQGPKEPFRD YVDRFFKTLR AEQASQEVKN WMTESLLVQN SNPDCKTILK
F2_CM_MP25  KQGPKEPFRD YVDRFFKTLR AEQATQEVKG WMTETLLVQN ANPDCKTILK
F2KU_BE_VI  KQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLIQN ANPDCKTILK
G_BE_DRCBL  RQGPKEPFRD YVDRFFKTLR AEQATQEVKS WMTDTLLIQN ANPDCKIILK
G_NG_92NG0  RQGPKEPFRD YVDRFFKTLR AEQATQEVKG WMTDTLLVQN ANPDCKTILR
G_SE_SE616  RQGPKEPFRD YVDRFFKCLR AEQASQDVKG WMTDTLLVQN ANPDCKTILR
H_BE_VI991  KQGPKEPFRD YVDRFFRVLR AEQATQDVKN WMTDTLLVQN ANPDCRTILK
H_BE_VI997  KQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILR
H_CF_90CF0  KQGPKEPFRD YVDRFFKTLR AEQATQDVKN WMTETLLVQN ANPDCKTILR
J_SE_SE702  RQGPKEPFRD YVDRFFKALR AEQATQDVKN WMTDTLLVQN ANPDCKTILK
J_SE_SE788  RQGPKEPFRD YVDRFFKALR AEQATQDVKN WMTDTLLVQN ANPDCKTILK
K_CD_EQTB1  RQGPKEPFRD YVDRFFRVLR AEQATQEVKN WMTETLLVQN ANPDCRTILK
K_CM_MP535  RQGPKEPFRD YVDRFFKTLR AEQATQEVKN WMTDTLLVQN ANPDCKTILK
N_CM_YBF30  KQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKQLLK
O_CM_ANT70  KQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKQILK
O_CM_MVP51  RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN SNPDCKQILK
O_SN_99SE_  RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKQILK
O_SN_99SE_  RQGPKEPFRD YVDRFYKTLR AEQATQEVKN WMTETLLVQN ANPDCKQILK
U_CD___83C  KQGPKEPFRD YVDRFFKVLR AEQSSQEVKN WMTDTLLIQN ANPDCKTILK

            351                                              400
00BW0762_1  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQANN... ...TNIMLQK
00BW0768_2  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQHN.... ...TSIMMQK
00BW0874_2  GLGPGASLEE MMTACQGVGG PSHKARVLAE AMSQSN.... ...S.IMMQR
00BW1471_2  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQANN... ...TNVMMQK
00BW1616_2  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQAGN... ...ATIMMQK
00BW1686_8  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQANS... ...PNILMQR
00BW1759_3  ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSKANN... ...ANILMQR
00BW1773_2  ALGPAASLEE MMTACQGVGG PSHKARVLAE AMSQANT... ...TNIMMQK
00BW1783_5  ALGQGASLEE MMTACQGVGG PGHKARVLAE AMSQVGN... ...PQVMMQR
00BW1795_6  ALGSGASLEE MMTACQGVGG PSHKARVLAE AMSQANN... ...VQIMMQK
00BW1811_3  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQTTS... ...MQIMMQR
00BW1859_5  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQANN... ...ANIMMQR
00BW1880_2  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQATN... ...TSIMMQK
00BW1921_1  ALGPGATLEE MMTACQGVGG PSHKARALAE AMSQANN... ...TNIMMQK
00BW2036_1  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVNG... ...SNVLMQR
00BW2063_6  ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSQANN... ...INVMMQK
00BW2087_2  ALGSGATLEE MMTACQGVGG PGHKARVLAE AMSQANS... ...TNIMIQR
00BW2127_2  ALGSGVTLEE MMSACRGVGG PSHKARVLAE AMSQANN... ...TNIMMQR
00BW2128_3  ALGPAATLVE MMTACQGVGG PSHKARVLAE AMSQTAS... ...TNILMQR
00BW2276_7  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQAN.... ...SNIMMQR
00BW3819_3  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQANN... ...INVMMQR
00BW3842_8  ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSQANS... ...TNIMMQR
00BW3871_3  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQANN... ...SNIMMQK
00BW3876_9  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQAGS... ...VNILMQR
00BW3886_8  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQANN... ...INVMMQR
00BW3891_6  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQTN.... ...SNIMMQN
00BW3970_2  AIGPGASLEE MMTACQGVGG PSHKARVLAE AMSQTH.... ...SNIMMQR
```

```
00BW5031_1  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQANN... ...INIMMQR
 96BW01B21  ALGPGASIEE MMTACQGVGG PSHKARVLAE AMSQANS... ...MNIMMQR
  96BW0407  ALGPGASLEE MMTACQGVGG PSHKARVLGE AMSQA.N... ...TNVMMQK
  96BW0502  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQANS... ...VNIMMQK
 96BW06_J4  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQANS... ...TSILMQR
 96BW11_06  TLGPGASLEE MMTACQGVGG PSHKARILAE AMSQANN... ...PNIMMQK
  96BW1210  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSHAGN... ...AGIMMQR
 96BW15B03  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQATS... ...ANILMQR
 96BW16_26  ASGPGASLEE MMTACQGVGG PGHKARVLAE AMSQANN... ...TNIMIQR
 96BW17A09  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSHANN... ...TSIMMQK
 96BWMO1_5  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQANN... ...IQVMMQK
 96BWMO3_2  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQAN.... ...TNIMMQR
98BWMC12_2  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSHTN.... ...S.IMMQR
98BWMC13_4  ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSQANN... ...INIMMQK
98BWMC14_a  SLGTGATLEE MMTACQGVGG PSHKARVLAE AMSQAN.... ...TSIMMQR
98BWMO14_1  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQANS... ...TNILIQR
98BWMO18_d  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQAGN... ...ANIMMQR
98BWMO36_a  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQANN... ...ANIMMQR
98BWMO37_d  ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSQAN.... ...SNIMMQR
99BW3932_1  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQANN... ...ASVMMQR
99BW4642_4  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMGQVNS... ...TNIMMQR
99BW4745_8  GLGPGATLEE MMTACQGVGG PSHKARVLAE AMSQTN.... ...SNIMMQR
99BW4754_7  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQANN... ...PNIMMQK
99BWMC16_8  ALGPGASIEE MMTACQGVGG PSHKARVSAE AMSQANQ... ...ANIMMQK
A2_CD_97CD  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVQNT.. ....NIMIQR
A2_CY_94CY  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSHVQSTN. ..T.NIMMQR
A2D___97KR  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQVQNTN. ..SNIMMMQK
A2G_CD_97C  .......... .......... .......... .......... ..........
A_BY_97BL0  ALGSEATLEE MMTACQGVGG PGHKAXVXAE AMSQVQN... ...ANIMMQK
A_KE_Q23_A  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVQH... ...TNIMMQR
A_SE_SE659  .......... .......... .......... .......... ..........
A_SE_SE725  .......... .......... .......... .......... ..........
A_SE_SE753  ALGAGATLEE MMTACQGVGG PGHKARVLAE AMSQVQH... ...PNIMMQR
A_SE_SE853  .......... .......... .......... .......... ..........
A_SE_SE889  .......... .......... .......... .......... ..........
A_SE_UGSE8  GLGAGATLEE MMTACQGVGG PSHKARVLAE AMSQAQQ... ...TNIMMQR
A_UG_92UG0  ALGAGATLEE MMTACQGVGG PGHKARVLAE AMSQVQH... ...TNIMMQR
A_UG_U455_  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVQQT.. ....SIMMQR
AC_IN_2130  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQTN.... ...NSILMQR
AC_RW_92RW  ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSQVQQ... ...PNIMMQR
AC_SE_SE94  .......... .......... .......... .......... ..........
ACD_SE_SE8  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQAN.... .ANTAIMMQR
ACG_BE_VI1  ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQANN... ....TVMMQR
AD_SE_SE69  ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQATNN.. .INAAIMMQR
AD_SE_SE71  .......... .......... .......... .......... ..........
ADHK_NO_97  ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVQN... ...VSVMMQR
ADK_CD_MAL  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQATN... .STAAIMMQR
AG_BE_VI11  ALGTGATLEE MVTACQGVGG PGHKARVLAE AMSQVH.... ..QTNIMMQR
AG_NG_92NG  ALGAGATLEE MLTACQGVGG PSHKARVLAE AMSRAT.G.. .TSAAIMMQK
AGHU_GA_VI  GLGAGATLEE MMTACQGVGG PGHKARVLAE AMSQVHN... ...TSIMMQR
AGU_CD_Z32  ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQAS.S.. .TAAAIMMQK
AJ_BW_BW21  ALGAGATLEE MMTACQGVG. PGHKAKILAE AMSQVQN... ...TNIMMQR
B_AU_VH_AF  ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSATIMMQR
B_CN_RL42_  ALGPAATLEE MMTACQGVGG PSHKARILAE AMSQVTNSAI TNSATIMMQR
B_DE_D31_U  ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSATVMMQR
B_DE_HAN_U  ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .GSAAIMMQR
B_FR_HXB2_  ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSATIMMQR
B_GA_OYI__  ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVN.... ..SVTVMMQK
B_GB_CAM1_  ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSATIMMQR
B_GB_GB8_A  ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVN.... ..SATIMMQK
B_GB_MANC_  ALGPAATLEE MMTACQGVGG PSHKARILAE AMSQVT.... .NSATIMMQR
```

```
B_KR_WK_AF   ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQAT.... .NSATIMMQR
B_NL_3202A   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSATIMMQR
B_TW_TWCYS   ALGPAATLEE MMTACQGVGG PSHKARVLAE AMSRVP.... .NSTNIMMQR
B_US_BC_L0   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSATIMMQR
B_US_DH123   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQITN... .TSATIMMQG
B_US_JRCSF   ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NPATIMMQR
B_US_MNCG_   ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSATIMMQR
B_US_P896_   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSATIMMQR
B_US_RF_M1   ALGPAATLEE MMTACQGVGG PSHKARILAE AMSQVT.... .NSATIMLQK
B_US_SF2_K   ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NPANIMMQR
B_US_WEAU1   ALGPGATLED MMTACQGVGG PGHKARVLAE AMSQVT.... .NTATMMMQR
B_US_WR27_   ALGPGATLEE MMTACQGVGG PGHKARVXAE AMSQVT.... .GSNAIMMQK
B_US_YU2_M   ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSATIMMQR
BF1_BR_93B   ALGPAATLEE MMTACQGVGG PGHKARVLAE AMSQVT.... .NSGTIMMQR
C_BR_92BR0   ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSKVNN... ...TNIMMQR
C_BW_96BW0   ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQT.N... ...TNVMMQR
C_BW_96BW1   ALGPGASLEE MMTACQGVGG PSHKARILAE AMSQANN... ...SNIMMQK
C_BW_96BW1   ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSHAGN... ...AGIMMQR
C_BW_96BW1   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQATS... ...ANILMQR
C_ET_ETH22   ALGPGASLEE MMTACQGVGG PAHKARVLAE AMSQVNN... ...TTIMMQK
C_IN_93IN1   ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSQAN.... ...STILMQR
C_IN_93IN9   ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQTN.... ...SAILMQK
C_IN_93IN9   ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSQAN.... ...SNILMQR
C_IN_94IN1   ALGPGASLEE MMTACQGVGG PSHKAKVLAE AMSQTN.... ...SAILMQR
C_IN_95IN2   ALGPGASLEE MMTACQGVGG PSHKARVLAE AMSQTN.... ...SAILMQR
CRF01_AE_C   ALGTGATLEE MMTACQGVGG PGHKARVLAE AMSQVQ.... ..HANIMMQR
CRF01_AE_C   ALGTGATLEE MMTACQGVGG PGHKARVLAE AMSQVH.... ..HTNIMMQK
CRF01_AE_C   ALGTGATLEE MMTACQGVGG PGHKARVLAE AMSQVQ.... ..HANIMMQR
CRF01_AE_T   SLGTGATLEE MMTACQGVGG PSHKARVLAE AMSHAQ.... ..HANIMMQR
CRF01_AE_T   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSQAQ.... ..HANIMMQR
CRF01_AE_T   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSQAQ.... ..HANIMMQR
CRF01_AE_T   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSHAQ.... ..QANIMMQR
CRF01_AE_T   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSHAQ.... ..HATIMMQR
CRF01_AE_T   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSHVQ.... ..QANIMMQR
CRF02_AG_F   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVQ.... ..QSNIMMQR
CRF02_AG_F   ALGPGASLEE MMTACQGVGG PGHKARVLAE AMSQVQ.... ..QSNIMMQR
CRF02_AG_G   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVQSP.. ....NILMQR
CRF02_AG_N   ALGTGATLEE MMTACQGVGG PGHKARVLAE AMSQVQ.... ..QANVMMQR
CRF02_AG_S   ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQVQ.... ..QPNIMMQR
CRF02_AG_S   ALGPGATLEE MMSACQGVGG PGHKARVLAE AMSQAQ.... ..QSNIMMQR
CRF03_AB_R   ALGSGATLEE MMTVCQGVGG PGHKARVLAE AMSQVQN... ...ANIMMQK
CRF03_AB_R   ALGSGATLEE MMTACQGVGG PGHKARVLAE AMSQVQN... ...ANIMMQK
CRF04_cpx_   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSQASN... .AAAAIMMQK
CRF04_cpx_   ALGTGATLEE MMTACQGVGG PAHKARVLAE AMSQASN... .AAAAIMMQK
CRF04_cpx_   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSQASS... .AAAAIMMQK
CRF05_DF_B   ALGQQATLEE MMTACQGVGG PSHKARVLAE AMSQATN... .SAATAMMQR
CRF05_DF_B   ALGPQATLEE MMTACQGVGG PSHKARVLAE AMSQATG... .SPAVMMQ.R
CRF06_cpx_   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQAS.... .VGAIMMQ.K
CRF06_cpx_   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQASGT.. .LTAIMMQ.K
CRF06_cpx_   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQASGT.. .VAAAIMMQK
CRF06_cpx_   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQVSGA.. .TAAIMMQ.K
CRF11_cpx_   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSQVQQ... ...TNVMMQR
CRF11_cpx_   ALGVGATLEE MMTACQGVGG PSHKARVLAE AMSQAQH... ...TNIMMQR
D_CD_84ZR0   ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQATS... .ASAAIMMQK
D_CD_ELI_K   ALGPQATLEE MMTACQGVGG PSHKARVLAE AMSQATN... .SVTTAMMQR
D_CD_NDK_M   ALGPQATLEE MMTACQGVGG PGHKARVLAE AMSQVTG... .SATAVMMQR
D_UG_94UG1   ALGPAATLEE MMTACQGVGG PSHKARVLAE AMSQATN... .ANTAIMMQR
F1_BE_VI85   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSQAN.... ...SAIMMQK
F1_BR_93BR   ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQATN... ...TAIMMQK
F1_FI_FIN9   GLGIGATLEE MMTACRGVGG PGHKARILAE AMSQAN.... ...TTIMMQK
F1_FR_MP41   ALGPGATLEE MMTACQGVGG PGHKARVLAE AMSQATN... ...AAIMMQK
```

```
F2_CM_MP25   ALGPGATLEE MMTACQGVGG PSHKARILAE AMSKATG... ...AAIMMQK
F2KU_BE_VI   ALGPGASLEE MMTACQGVGG PAHKARVLAE AMSQATN... ...TAIMMQK
G_BE_DRCBL   GLGQGATLEE MMTACQGVGG PSHKARVLAE AMSQAS.G.. .AAAAIMMQK
G_NG_92NG0   ALGPGATLEE MMTACQGVGG PSHKARVLAE AMSQASGA.. .AAAAIMMQK
G_SE_SE616   ALGQGASLEE MMTACQGVGG PSHKARVLAE AMSQAS.G.. .AAAAIMMQR
H_BE_VI991   ALGRGASIEE MMTACQGVGG PSHKARVLAE AMSQVTNAS. ...AAIMMQK
H_BE_VI997   ALGQGASIEE MMTACQGVGG PSHKARVLAE AMSQVTNAN. ...AAIMMQK
H_CF_90CF0   ALGQGASIEE MMTACQGVGG PSHKARVLAE AMSQVTNTN. ...TAIMMQK
J_SE_SE702   ALGSGATLEE MMTACQGVGG PGHKARVLAE AMSQVTN... ...TNIMMQR
J_SE_SE788   ALGSGATLEE MMTACQGVGG PGHKARVLAE AMSQVTN... ...TNIMMQR
K_CD_EQTB1   ALGSGATLEE MMTACQGVGG PGHKARVLAE AMSQVTN... ...SAVMMQR
K_CM_MP535   ALGPGASLEE MMTACQGVGG PSHKARILAE AMSQVTN... ...PVVMMQK
N_CM_YBF30   ALGPGATLEE MMTACQGVGG PAHKARVLAE AMSQVQQP.. ..TTSVFAQR
O_CM_ANT70   SLGPGATLEE MMVACQGVGG PTHKARVLAE AMATAQQDLK GGYTAVFMQR
O_CM_MVP51   ALGPEATLEE MMVACQGVGG PTHKAKILAE AMASAQQDLK GGYTAVFMQR
O_SN_99SE_   SLGPGATLEE MMIACQGVGG PTHKARVLAE AMA.AAQDLK GGYTAVFMQR
O_SN_99SE_   SLGPGATLEE MMIACQGVGG PTHKARVLAE AMS.AAQDLK GGYSAVFMQR
U_CD___83C   ALGTGATLEE MMTACQGVGG PSHKARVLAE AMSQTN.... ...TAIMMQR

             401                                                 450
00BW0762_1   NNFKGPRRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW0768_2   SNFKGPKRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW0874_2   SNFKGHKRIV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW1471_2   GNFKGPRRVI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKECT..E
00BW1616_2   SNFKGPRRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW1686_8   SNFKGSK.RV KCFNCGKEGH IARNCRAPRK RGCWKCGKEG HQMKDCT..E
00BW1759_3   SNFKGPKRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW1773_2   SNFKGPRRTV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW1783_5   NNLKGTRRTV KCFNCGKEGH IARNCRAPRK KGCWKCGREG HQMKDCT..E
00BW1795_6   NNFKGPRRII KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW1811_3   SNFKGSKRSV KCFNCGKEGH IARNCRAPRK KGCWKCGKDG HQMKDCT..E
00BW1859_5   SNFKGPRKII KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW1880_2   SNFKGPRRIV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW1921_1   SNFKGPRRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW2036_1   GNFKGPKRII KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW2063_6   NNFKGPRRLV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW2087_2   SNFKGSKRIV KCFNCGKEGH IAKNCRAPRK KGCWKCGREG HQMKDCT..E
00BW2127_2   GNFRGAKRSV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW2128_3   SNFKGSKRSV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW2276_7   SNFKGPRKIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCTN.D
00BW3819_3   SNFKGSKRIV KCFNCGKEGH IARNCRAPRK KGCWKCGREG HQMKDYT..E
00BW3842_8   GNFKGPKRTI KCFNCGKEGH LARNCRAPKK KGCWKCGREG HQMKDCT..E
00BW3871_3   SNFKGPRRII KCFNCGKEGH LARNCRAPRK RGCWKCGKEG HQMKDCTT.E
00BW3876_9   GNFKGSKRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW3886_8   GNFKGAKRIV KCFNCGREGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW3891_6   SNFKGSRRIV KCFNCGKVGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
00BW3970_2   SNFKGPKRIV KCFNCGKEGH IAKNCRAPRK KGCWKCGQEG HQMKDCT..E
00BW5031_1   NNFKGPKRIV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
 96BW01B21   SNFKNPKRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
  96BW0407   SNFKGPRRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
  96BW0502   SNFKGPRRNV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
 96BW06_J4   GNFKGPKRII KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
 96BW11_06   NNFKGPRRIV KCFNCGKEGH IARNCKAPRK KGCWKCGKEG HQMKDCT..E
  96BW1210   GNFKGPRKIP KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCS..E
 96BW15B03   SNFKGPKRII KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
 96BW16_26   SNFKGPRRSV KCFNCGKEGH IARNCRAPRK KGCWKCGQEG HQMKDCT..E
 96BW17A09   GNFKGPRRTI KCFNCGKEGH LAKNCRAPRK XGCWKCGKEG HQMKECT..E
 96BWMO1_5   NNFKGPRRTI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
 96BWMO3_2   GNFKGPKRII KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCNN.E
98BWMC12_2   SNFKGPKRIV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCN..E
98BWMC13_4   SNFKGPKRII KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
```

```
98BWMC14_a  GNFKGPRRII KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
98BWMO14_1  SHFKGSKRTV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
98BWMO18_d  GNFKGPRRII KCSNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
98BWMO36_a  SNFKGPKRTV KCFNCGKEGH VARNCRAPRK KGCWKCGREG HQMKDCT..E
98BWMO37_d  GNFKGSKRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
99BW3932_1  SNFKGPKRII KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
99BW4642_4  SNFKGPKRII KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
99BW4745_8  SNFKGPRRTV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
99BW4754_7  XNFKGPRRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
99BWMC16_8  SNFKGPKRIV KCFNCGKEGH IARNCRAPRK KGCWKCGREG HQMKDCT..E
A2_CD_97CD  GNFKGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
A2_CY_94CY  GNFRGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
A2D___97KR  GNFRGQK..I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
A2G_CD_97C  .......... .......... .......... .......... ..........
A_BY_97BL0  SNFRGPKR.I KCXNCGKEGH LARNCRAPRK XGCWKCGKEG HQMKDCTE..
A_KE_Q23_A  GNFKGQKR.I KCFNCGKEGH LARNCRALRK KGCWKCGKEG HQMKDCTE..
A_SE_SE659  .......... .......... .......... .......... ..........
A_SE_SE725  .......... .......... .......... .......... ..........
A_SE_SE753  GNFRGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
A_SE_SE853  .......... .......... .......... .......... ..........
A_SE_SE889  .......... .......... .......... .......... ..........
A_SE_UGSE8  GNFRGQKK.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
A_UG_92UG0  GNFKGQKR.I KCFNCGKEGH LAKNCRAPRK KGCWKCGREG HQMKDCTE..
A_UG_U455_  GNFRGPRR.I KCFNCGKEGH LAKNCRAPRK KGCWKCGKEG HQMKDCT..E
AC_IN_2130  SNFKGFKRTV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
AC_RW_92RW  GNFKGQRRII KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
AC_SE_SE94  .......... .......... .......... .......... ..........
ACD_SE_SE8  GNFKGPKKII KCFNCGKEGH IAKNCRAPRK KGCWKCGREG HQMKDCTE..
ACG_BE_VI1  GNFKGPKRII KCFNCGKEGH VARNCRAPRK KGCWKCGKEG HQMKDCTT.E
AD_SE_SE69  GNFKGQRKII KCFNCGKLGH IARNCKAPRK KGCWKCGKEG HQMKDCTE..
AD_SE_SE71  .......... .......... .......... .......... ..........
ADHK_NO_97  GNFKGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
ADK_CD_MAL  GNFKGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
AG_BE_VI11  GNFRGQKT.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
AG_NG_92NG  NNFKGPRRGI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
AGHU_GA_VI  GNFKGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
AGU_CD_Z32  SNFKGPRKTI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
AJ_BW_BW21  SNFKGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDC.E..
B_AU_VH_AF  GNFRNQRKTV KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
B_CN_RL42_  GNFRNQRKIV KCFNCGKEGH IARNCRAPRK RGCWKCGKEG HQMKDCTE..
B_DE_D31_U  GNFRNQRKTV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HLMKDCTE..
B_DE_HAN_U  GNFRNQRKTV KCFNCGKEGH IAKNCRAPRR KGCWKCGKEG HQMKDCTE..
B_FR_HXB2_  GNFRNQRKIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCTE..
B_GA_OYI__  GNFKNQRKTV KCFNCGKEGH TARNCRAPRK KGCWKCGKEG HQMKDCTE..
B_GB_CAM1_  GNFRNQRKTV KCFNCGKEGH IAKNCRAPRK KGCWKCGREG HQMKDCTE..
B_GB_GB8_A  GNFRSQRKTV KCFNCGKVGH IAKNCRAPRK KGCWKCGKEG HQMKDCNE..
B_GB_MANC_  GNFRNQRKTV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKECTE..
B_KR_WK_AF  GNFRNQRRTV KCFNCGKEGH IAKNCRAPRR KGCWKCGKEG HQMKDCTE..
B_NL_3202A  GNFRNQRKIV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCTE..
B_TW_TWCYS  GNYRNQRKTV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCTE..
B_US_BC_L0  GNFRNQRKTV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCTE..
B_US_DH123  GNFRNQRK.I KCFNCGKEGH IARNCKAPRK RGCWKCGKEG HQMKDCTE..
B_US_JRCSF  GNFRNQRKNV KCFNCGKEGH ISKNCRAPRK KGCWKCGKEG HQMKDCTE..
B_US_MNCG_  GNFRNQRKII KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKECTE..
B_US_P896_  GNFRNQRKTV KCFNCGKEGH IAKNCRAPRK RGCWKCGKEG HQMKDCTE..
B_US_RF_M1  GNFRDQRKIV KCFNCGKVGH IAKNCRAPRK KGCWKCGKEG HQMKDCTE..
B_US_SF2_K  GNFRNQRKTV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCTNEG
B_US_WEAU1  GNFRSPRKTI KCFNCGKEGH IARNCRAPRK KGCWKCGQEG HQMKDCTE..
B_US_WR27_  GNFRNQRKTV KCXXCGKEGH XARXCKAPRK RGCWKCGKEG HQMXDXXE..
B_US_YU2_M  GNFRNQRKTV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCTE..
BF1_BR_93B  GNFRNQRKTI KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCTE..
```

```
C_BR_92BR0   SNCKGPKRTI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQVKDCT..E
C_BW_96BW0   SNFKGPRRIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
C_BW_96BW1   NNFKGPTRIV KCFNCGKEGH IARNCKAPRK KGCWKCGKEG HQMKDCT..E
C_BW_96BW1   GNFKGPRKIP KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCS..E
C_BW_96BW1   SNFKGPKRII KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
C_ET_ETH22   SNFKGPKRAI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
C_IN_93IN1   SNFKGSKRIV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
C_IN_93IN9   SNFKGSKRII KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
C_IN_93IN9   SNFKGSKRTV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
C_IN_94IN1   SNFKGSKRII KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
C_IN_95IN2   SNFKGSKRIV KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
CRF01_AE_C   GNFKGQKR.I KCFNCGKEGH LARNCRAPRK RGCWKCGQEG HQMKDCTE..
CRF01_AE_C   GNFKGQRK.I KCFNCGKEGH LARNCRAPRK KGCWKCGREG HQMKDCTE..
CRF01_AE_C   GNFKGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF01_AE_T   GNFKGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF01_AE_T   GNFKGPRR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF01_AE_T   GNFKGQTR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF01_AE_T   GNFKGQKR.I KCFNCGKEGH LARNCRALRK KGCWKCGKEG HQMKDCTE..
CRF01_AE_T   GNFKGQKR.I KCFNCGREGH LARNCRAPRK QGCWKCGKEG HQMKDCTE..
CRF01_AE_T   GNFKGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF02_AG_F   GNFRGQRT.I KCFNCGKEGH LARNCKAPRK KGCWKCGKEG HQMKDCTE..
CRF02_AG_F   GNFRGQRT.I KCFNCGKEGH LARNCKAPRK KGCWKCGKEG HQMKDCTE..
CRF02_AG_G   GNFKGQK.RI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCT..E
CRF02_AG_N   GNFRGQRT.I KCFNCGKEGH LARNCKAPRK RGCWKCGKEG HQMKDCTE..
CRF02_AG_S   GNFRGQRP.I KCFNCGKEGH LARNCKAPRK KGCWKCGKEG HQMKDCTE..
CRF02_AG_S   GNFRGQRT.I KCFNCGKEGH LARNCKAPRK RGCWKCGKEG HQMKDCTE..
CRF03_AB_R   SNFRGPKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF03_AB_R   SNFRGPKR.I KCFNCGKDGH LARNCRAPRK KGCWKCGKEG HQMKDCNE..
CRF04_cpx_   SKFKGQRRTI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF04_cpx_   SNFRGQKRII KCFNCGKEGH LARNCRAPRK RGRWKCGKEG HQMKDCTE..
CRF04_cpx_   SNFKGPRRSI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCPE..
CRF05_DF_B   GNFKGPRKII KCFNCGKEGH IAKNCRAPGK KGCWKCGREG HQMKDCTE..
CRF05_DF_B   GNFKGPRKSI KCFNCGKEGH TAKNCRAPRK RGCWKCGREG HQMKDCIE..
CRF06_cpx_   SNFKGPKRSI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF06_cpx_   SNYKGPKRSI KCFNCGREGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF06_cpx_   SNFKGP.RKI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF06_cpx_   SNFKGQRKNI KCFNCGKEGH TARNCRAPRK KGCWKCGKEG HQMKDCTE..
CRF11_cpx_   SNFKGQKR.I KCFNCGKEGH LARNCRAPRK KGCWKCGREG HQMKDCTE..
CRF11_cpx_   GNFKGQKR.I KCFNCGKEGH LARNCRAPRK K.GCKCGKEG HQMKDCTE..
D_CD_84ZR0   SNFKGTRKIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCTE..
D_CD_ELI_K   GNFKGPRKII KCFNCGKEGH IAKNCRAPRK KGCWRCGKEG HQLKDCTE..
D_CD_NDK_M   GNFKGPRKSI KCFNCGKEGH TAKNCRAPRK KGCWKCGREG HQMKDCTE..
D_UG_94UG1   GNFKGPKKII KCFNCGKEGH TAKNCRAPRK KGCWKCGREG HQMKDCTE..
F1_BE_VI85   SNFKGQRRVV KCFNCGKEGH IARNCRAPRK KGCWKCGREG HQMKDCT..E
F1_BR_93BR   SNFKGQRRIV KCFNCGKEGH IAKNCRAPRK KGCWKCGREG HQMKDCT..E
F1_FI_FIN9   SNFRGQRRIV KCFNCGKEGH IARNCRAPRK KGCWKCGQEG HQMKDCT..E
F1_FR_MP41   SNYKGPRRFI KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
F2_CM_MP25   SNFKGQRRIV KCFNCGKEGH IARNCRAPRK RGCWKCGQEG HQMKDCT..E
F2KU_BE_VI   GNFKGPRRDV KCFNCGKEGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
G_BE_DRCBL   SNFKGPRRTI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKECTE..
G_NG_92NG0   SNFKGPRRII KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKECTE..
G_SE_SE616   SNFKGPRRTI KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCTE..
H_BE_VI991   GNFKGPRRTV KCSNCGKEGH IARNCRAPRK KGCWKCGQEG HQMKDCT..G
H_BE_VI997   SNFKGPRKIV KCFNCGKEGH IARNCRAPRK KGCWKCGREG HQMKDCT..E
H_CF_90CF0   GNFKGQRKFV KCFNCGKEGH IARNCRAPRK KGCWKCGREG HQMKDCT..E
J_SE_SE702   GNFRDHKRIV KCFNCGKQGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
J_SE_SE788   GNFRDHKRIV KCFNCGKQGH IAKNCRAPRK KGCWKCGKEG HQMKDCT..E
K_CD_EQTB1   GNFKGQRRII KCFNCGKEGH LARNCRAPRK KGCWKCGKEG HQMKDCS..E
K_CM_MP535   GNFKGHRKIV KCFNCGKEGH IARNCRAPRK KGCWKCGKEG HQMKDCT..E
N_CM_YBF30   GNFKGIRKPI KCFNCGKEGH LARNCKAPRR GGCWKCGQEG HQMKDCKNEG
O_CM_ANT70   GQNPIRKGTI KCFNCGKEGH IARNCRAPRK KGCWKCGQEG HQMKDCRN.G
```

```
O_CM_MVP51  GQNPNRKGPI KCFNCGKEGH IAKNCRAPRK RGCWKCGQEG HQMKDCKN.G
O_SN_99SE_  GQNPSRKGPI KCFNCGKEGH LARNCRAPRK KGCWKCGQEG HQMKDCKN.G
O_SN_99SE_  GQNPGRKGPI KCFNCGKEGH LARNCRAPRK KGCWKCGQEG HQMKDCRN.G
U_CD___83C  GNFKGPRRIV KCFNCGKEGH IAKNCRAPRK KGCWKCGREG HQMKDCT..E

                451                                                 500
00BW0762_1  RQANFLGKIW PSHKG.RPGN FLQSR..... .......PEP TAPPAESFK.
00BW0768_2  RQANFLGKIW PSHKG.RPGN FLQNRPEP.. .......... TAPPAESFK.
00BW0874_2  RQANFLGKIW PSHKG.RPGN FLQNRPEPSA PPAESLRPEP SAPPAESLR.
00BW1471_2  RQANFLGKIW PSQKG.RPGN FLQNRPEP.. .......... SAPPAESFR.
00BW1616_2  RQANFLGKIW PSHKG.RPGN FLQSRPEPTA PP....APVP TAPPAESFR.
00BW1686_8  RQANFLGKIW PSHKG.RPGN FLQNRPEP.. .......... SAPPAESFK.
00BW1759_3  RQANFLGKIW PSHKG.RPGN FLQ..... .....SRPEP TAPPLESFK.
00BW1773_2  RQANFLGKIW PSHKG.RPGN FLQSRPE... .........P TAPPAESFR.
00BW1783_5  RQANFLGKIW PSQKGGRPGN FLQNRPA... ....ESRLEP TAPPAESFR.
00BW1795_6  RQANFLGKIW PSHKG.RPGN FLQNRPE... .........P TAPPAESFR.
00BW1811_3  RQANFLGRIW PSHKG.RPGN FLQNRPEPTA P......LEP TAPPAESFR.
00BW1859_5  RQANFLGKIW PSHKG.RPGN FLQNRPEP.. .......... TAPPAESFR.
00BW1880_2  RQANFLGKIW PSHKG.RPGN FLQSR..... .......PEP TAPPAESFK.
00BW1921_1  RQANFLGKIW PSHKG.RPGN FLQSRPEP.. .......... TAPPAESFR.
00BW2036_1  RQANFLGKIW PSNKG.RPGN FLQNRTAPPV ........EP TAPPAESFR.
00BW2063_6  RQANFLGKIW PSHKG.RPGN FLQSRLE... .........P TAPPAESF..
00BW2087_2  RQANFLGKIW PSHKGGRPGN FLQSRPEPTA P.....PAEP TAPPAESFR.
00BW2127_2  RQANFLGKIW PSHKG.RPGN FLQNRPEPTA P.....RPEP SAPPAESFR.
00BW2128_3  RQANFLGRIW PSNKG.RPGN FLQNRPEPTA PPAE.NRPEP TAPPAESFR.
00BW2276_7  RQANFLGKLW PSNKG.RPGN FLQNRTEPTA P......LEP TAPPADSFK.
00BW3819_3  RQANFLGKIW PSHKG.RPGN FLQNRPE... P......TAP TAPPAESFR.
00BW3842_8  RQANFLGKIW PSRGG.RPGN FLQNRTEPTA P......PEP TAPPAESFR.
00BW3871_3  RQANFLGKIW PSHKG.RPGN FLQNRPEP.. .......... TAPPAESFR.
00BW3876_9  RQANFLGKIW PSHKG.RPGN FLQNRPE... .........P TAPPAESFR.
00BW3886_8  RQANFLGKIW PSHKG.RPGN FLQNRPEPTA P.....PAEP TAPPAESFR.
00BW3891_6  RQANFLGRIW PSHKG.RPGN FLQSRPE... .........P TAPPAESFR.
00BW3970_2  RQANFLGRFW PSQKG.RPGN FLQ..... .....SRSEP TAPPAESFR.
00BW5031_1  RQANFLGKIW PSNKG.RPGN FLQSRPEPTA P......PMP TAPPAESFR.
 96BW01B21  RQANFLGKIW PSHKG.RPGN FLQNR..... .......LEP SAPPAESFR.
  96BW0407  RQANFLGKIW PSHKG.RPGN FLQ..... .....SRPEP TAPPAESFR.
  96BW0502  RQANFLGKIW PSHKG.RPGN FLQNRSEPA. .....APTVP TAPPAESFR.
 96BW06_J4  RQANFLGKIW PSHKGGRPGN FLQSRPEP.. .......... TAPPAESFR.
 96BW11_06  RQANFLGKIW PSHKG.RPGN FLQSRPE... .........P TAPPAESL..
  96BW1210  GQANFLGKIW PSHKG.RPGN FLQSR..... .......PEP SAPPAESFR.
 96BW15B03  RQANFLGKIW PSHKG.RPGN FLQNRTEP.. .......... TAPPAESFK.
 96BW16_26  RQADFLGKIW PSHKG.RPGN FLQSRPE... .........P TAPPAESFR.
 96BW17A09  RQANFLGKIW PSHKGGRPGN FLQNRPEP.. .......... TAPPAESFR.
 96BWMO1_5  RQANFLGKIW PSHKG.RPGN FLQSRPE... .........P TAPPAEIL..
 96BWMO3_2  RQANFLGKIW PSHKG.RPGN FLQSRPEP.. .......... TAPPAERFR.
98BWMC12_2  RQANFLGRLW PSHKG.RPGN FPQNR..... .......VEP TAPPAESLR.
98BWMC13_4  KQANFLGKIW PSHKG.RPGN FLQSRPE... .........P TAPPAESL..
98BWMC14_a  RQANFLGKIW PSHKG.RPGN FLQSRPEP.. .......... SAPPAESFR.
98BWMO14_1  RQANFLGKIW PSHKGGRPGN FLQRRPEP.. .......... TAPPAESFR.
98BWMO18_d  RQANFLGKIW PSHKG.RPGN FIQNRPAPT. ....APPVEP TAPPAESFR.
98BWMO36_a  RQANFLGRIW PSHKG.RPGN FLQSRPEPTA P.....PAEP TAPPAESFR.
98BWMO37_d  RQANFLGKIW PSHKG.RPGN FLQ..... .....KRPEP TAPPAESFR.
99BW3932_1  RQVNFLGKIW PSNKG.RPGN FLQNRTVPTA PPAESFRIEP TAPPAESFR.
99BW4642_4  RQANFLGKIW PSHKG.RPGN FFQNRTEP.. .......... TAPPAESFR.
99BW4745_8  RQANFLGKIW PSNKG.RPGN FLQNRPEPTA P......LEP TAPPAESFR.
99BW4754_7  RQANFLGKIW PSNKG.RPGN FLQSR..... .......PEP TAPPAESFK.
99BWMC16_8  RQANFLGKIW PSNKG.RPGN FLQNRPEPT. .....APLEP TAPPAESFR.
A2_CD_97CD  RQANFLGKIW PSNKG.RPGN FPQSRTE... .........P TAPPME....
A2_CY_94CY  RQANFLGKIW PSNKG.RPGN FPQSRTE... .........P TAPPAENLR.
A2D___97KR  RQANFLGKIW PSHSG.RPGN FPQSRTE... .........P TAPPAEDFG.
A2G_CD_97C  .......... .......... .......... .......... ..........
```

```
A_BY_97BL0   RQANFLGRIW PSSKG.RPGN FPQSRPE... .........PS APP.AENFR.
A_KE_Q23_A   RQANFLGKIW PSRKG.RPGN FPQNRLE... .........PT APP.AETCG.
A_SE_SE659   .......... .......... .......... .......... ..........
A_SE_SE725   .......... .......... .......... .......... ..........
A_SE_SE753   RQANFLGRIW PSSKG.RPGN FPQSRLE... .........PT APP.AEIFG.
A_SE_SE853   .......... .......... .......... .......... ..........
A_SE_SE889   .......... .......... .......... .......... ..........
A_SE_UGSE8   RQANFLGKIW PSHKG.RPGN FPQSRPE... .........PS APP...AEM.
A_UG_92UG0   RQANFLGKIW PSSKG.RPGN FPQSRPE... .........PT APPAAEIFG.
A_UG_U455_   RQANFLGKIW PSNKG.RPGN FPQSRPE... ..........P TAPPAEIFG.
AC_IN_2130   RQANFLGKIW PSHKG.RPGN FLQ....... .....NRPEP TAPPAESFR.
AC_RW_92RW   RQANFLGKIW PSNKG.RPGN FPQSRL.... .........EP TAPPA.....
AC_SE_SE94   .......... .......... .......... .......... ..........
ACD_SE_SE8   RQANFLGKIW PSHKG.RPGN FLQSRPE... .........PT APPA.ESFG.
ACG_BE_VI1   RQANFLGKIW PSNKG.RPGN FPQSRPEPTA PP.....AEP TAPPAESFG.
AD_SE_SE69   RQANFLGKIW PSSKG.RPGN FLQSRP.... .........EP TAPPAESFG.
AD_SE_SE71   .......... .......... .......... .......... ..........
ADHK_NO_97   RQANFLGKIW PSSKG.RPGN FPQSRPE... .........PS APPA.ESFG.
ADK_CD_MAL   RQANFLGKIW PSHKG.RPGN FLQSRPE... .........PT APPA.ESFG.
AG_BE_VI11   RQANFLGKIW PSSKG.RPGN FPQSRLE... .........PT APPA.ESLG.
AG_NG_92NG   RQANFLGKIW PSNKG.RPGN FLQNRPE... ..........P TAPPAESFG.
AGHU_GA_VI   RQANFLGKIW PSNKG.RPGN FLQNRPE... .........PT APPA.ESFG.
AGU_CD_Z32   RQANFLGKIW PSNKG.RPGN FLQNRPE... ..........P TAPPAESFE.
AJ_BW_BW21   RQANFLGKIW PSNKG.RPGN FLQSRPE... .........PT APPA.ESFG.
B_AU_VH_AF   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_CN_RL42_   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_DE_D31_U   RQANFLGKIW PSYKG.RPGN FLQRRPE... ..........P TAPPEESFR.
B_DE_HAN_U   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_FR_HXB2_   RQANFLGKIW PSYKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_GA_OYI__   RQANFLGKIW PSHKG.RPGN FLQNRPE... ..........P TAPPAESFG.
B_GB_CAM1_   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_GB_GB8_A   RQANFLGKIW PSHKG.RPGN FLQSRPEPIA PP......EP TAPPEESFR.
B_GB_MANC_   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_KR_WK_AF   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P SAPPEESFR.
B_NL_3202A   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_TW_TWCYS   RQANFLGKIW PSHKE.RPGN FLQSRPE... ..........P TAPPEESFR.
B_US_BC_L0   RQANFLGKIW PSHKG.RPGN FPQSRLE... ..........P TAPPEESFR.
B_US_DH123   RQANFLGKIW PSHKE.RPGN FLQSRPE... ..........P SAPPEESFR.
B_US_JRCSF   RQANFLGKIW PSYKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_US_MNCG_   RQANFLGKIW PSCKG.R.RN FPQSRTE... ..........P TAPPEESFR.
B_US_P896_   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_US_RF_M1   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_US_SF2_K   RQANFLGKIW PSYKG.RPGN FLQSRPE... ..........P TAPPEESFR.
B_US_WEAU1   RQANFLGKIW SSQKG.RPGN FPQSRLE... ..........P TAPPEESFR.
B_US_WR27_   RQAXFLGXIR PSHXG.RPGX FLQNRPE... ..........P SAPPAESFR.
B_US_YU2_M   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPSEESVR.
BF1_BR_93B   RQANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPAESFR.
C_BR_92BR0   RQANFLGKIW PSHRG.RPGN LLQNRT.... .........EP TAPPE.....
C_BW_96BW0   RQANFLGKIW PSHKG.RPGN FLQ....... .....SRPEP TAPPAESFR.
C_BW_96BW1   RRANFLGKIW PSHKG.RPGN FLQSRPE... ..........P TAPPAESF..
C_BW_96BW1   GQANFLGKIW PSHKG.RPGN FLQSR..... .......PEP SAPPAESFR.
C_BW_96BW1   RQANFLGKIW PSHKG.RPGN FLQNRTEP.. .......... TAPPAESFK.
C_ET_ETH22   RQANFLGRLW PSNKG.RPGN FLQSRP.... .........EP TAPPESLRPE
C_IN_93IN1   RQANFLGKIW PSHKG.RPGN FLQ....... .....SRPEP TAPPAESFR.
C_IN_93IN9   RQANFLGKIW PSHKG.RPGN FLQ....... .....SRPEP TAPPAESFR.
C_IN_93IN9   RQANFLGKIW PSHKG.RPGN FLQNRPEPTA PP...ARPEP TAPPAESFR.
C_IN_94IN1   RQANFLGKIW PSHKG.RPGN FLQ....... .....SRPEP TAPPAESFR.
C_IN_95IN2   RQANFLGKIW PSHKG.RPGN FLQ....... .....SRPEP TAPPAESFR.
CRF01_AE_C   RQANFLGKIW PLNKG.RPGN FPQSRLE... .........PT APPA.ESLG.
CRF01_AE_C   RQANFLGKIW PSSKG.RPGN FPQSRPE... .........PT APPM.ESLG.
CRF01_AE_C   RQANFLGRIW PSSKG.RPGN FPQSRPE... .........PT APPA.ESLG.
```

```
CRF01_AE_T  RQANFLGKFW PSNKG.RPGN FPQSRPE... ........PT APPA.ENWG.
CRF01_AE_T  RQANFLGKIW PSNKG.RPGN FPQSRPE... ........PT APP..AEWG.
CRF01_AE_T  RQANFLGKIW PSNKG.RPGN FPQSKPE... ........PT APPA.ENWG.
CRF01_AE_T  RQANFLGKIW PSNKG.RPGN FPQSRPE... ........PT APPA.ENWG.
CRF01_AE_T  RQANFLGKIW PSNKG.RPGN FPQSRPE... ........PT APPA.ENWG.
CRF01_AE_T  RQANFLGKIW PSNKG.RPGN FPQSRPE... ........PT APPA.ENWG.
CRF02_AG_F  GQANFLGKIW PSSKG.RPGN FPQSRPE... ........PT APPA.ESLG.
CRF02_AG_F  RQANFLGKIW PSSKG.RPGN FPQSRPE... ........PT APPA.ESFG.
CRF02_AG_G  RQANFLGKIW PSNKG.RPGN FPQSRPE... P......... SAPPAESFG.
CRF02_AG_N  RQANFLGKIW PSSKG.RPGN FPQSRPE... ........PT APPA.ESFG.
CRF02_AG_S  RQANFLGKIW PSSKG.RPGN FPQSRPE... ........PT APPA.ESLG.
CRF02_AG_S  RQANFLGKIW PSSKG.RPGN FPQSRPE... ........PT APPA.ESFG.
CRF03_AB_R  RQANFLGRIW PSSKG.RPGN FPQSRPE... ........PS APP.AENFG.
CRF03_AB_R  RQANFLGKIW PSSKG.RPGN FPQSRPE... ........PS APP.AENFG.
CRF04_cpx_  RQANFLGRMW PSSKG.RPGN FLQNRPE... ........PT APPA.ECLE.
CRF04_cpx_  RQANFLGRMW PSSKG.RPGN FLQSRPE... ........PT APPA.ESLE.
CRF04_cpx_  RQANSLGRMW PSSKG.RPGN FLQSRTE... ........PT APPA.ESFE.
CRF05_DF_B  RQANFLGKVW PSHKG.RPGN FLQSRP.... ........EP SAPPAESFR.
CRF05_DF_B  GQANFLGRVW LSHKG.RPGN FLQSRP.... ........EP SAPPAESFG.
CRF06_cpx_  RQANFLGKIW PSNKG.RPGN FLQNRPE... .........P TAPPIESFG.
CRF06_cpx_  RQANFLGKIW PSNKG.RPGN FLQNRPE... .........P TAPPAESFG.
CRF06_cpx_  RQANFLGRIW PSSKG.RPGN FLQNRPE... .........P TAPPAESFG.
CRF06_cpx_  RQANFLGKIW PSHKG.RPGN FLQNRPEQNR P.......EP SAPPAESFG.
CRF11_cpx_  RQANFLGKIW PSSKG.RPGN FLQSRPE... ........PT APPA.ESFG.
CRF11_cpx_  RQANFLGKIW PSSKG.RPGN FLQSRPE... ........PT APPA.ESFG.
D_CD_84ZR0  RQANFLGKIW PSHKG.RPGN FLQSRPE... .........P TAPPAE.FG.
D_CD_ELI_K  RQANFLGRIW PSHKG.RPGN FLQSRP.... ........EP TAPPAESFG.
D_CD_NDK_M  RQANFLGKIW PSHKG.RPGN FLQSRP.... ........EP TAPPAESFG.
D_UG_94UG1  RQANFLGKIW PSHNG.RPGN FLQSRPPA.. ........EP TAPPAEIFG.
F1_BE_VI85  RQANFLGKIW PSNKG.RPGN FLQSRPE... .........P TAPPAESFG.
F1_BR_93BR  RQANFLGKIW PSNKG.RPGN FIQNRPE... .........P SAPPAESFR.
F1_FI_FIN9  RQANFLGKIW PSNKG.RPGN FLQSRPE... .........P TAPPAESLG.
F1_FR_MP41  RQANFLGKIW PSNKG.RPGN FLQNRPE... .........P TAPPAESFG.
F2_CM_MP25  RQANFLGKMW PSNKG.RPGN FLQNRPE... .........P TAPPAESFG.
F2KU_BE_VI  RQANFLGKIW PSNKG.RPGN FLQSRPE... .........P TAPPAESFG.
G_BE_DRCBL  RQANFLGKIW PSNKG.RPGN FLQNRPE... .........P TAPPAENFG.
G_NG_92NG0  RQANFLGKIW PSNKG.RPGN FLQNRTE... .........P TAPPAESFG.
G_SE_SE616  RQANFLGKIW PSNKG.RPGN FLQNRTE... .........P TAPPAESLG.
H_BE_VI991  RQANFLGKIW PSSKG.RPGN FPQKRLE... .........P TAPPAESFG.
H_BE_VI997  RQANFLGKIW PSSKG.RPGN FLQSRPE... .........P TAPPAESFG.
H_CF_90CF0  RQANFLGKIW PSSKG.RPGN FLQSRPE... .........P TAPPAESFG.
J_SE_SE702  RQANFLGKIW PSSKG.RPGN FLQSRPE... .........P TAPPAESLG.
J_SE_SE788  RQANFLGKIW PSSKG.RPGN FLQSRPE... .........P TAPPAESLG.
K_CD_EQTB1  RQANFLGKFW PLNKE.RPGN FLQNRPE... .........P TAPPAESFG.
K_CM_MP535  RQANFLGKIW PSHKG.RPGN FLQSRPE... .........P TAPPAESFG.
N_CM_YBF30  RQANFLGKSW SPFKG.RPGN FPQTTTRK.. ........EP TAPPLESYG.
O_CM_ANT70  KQANFLGKYW PP.GGTRPGN YVQRPAH... .........P SAPPMEEEVK
O_CM_MVP51  RQANFLGKYW PP.GGTRPGN YVQKQVS... .........P SAPPMEEAVK
O_SN_99SE_  RQANFLGKYW PP.GGTRPGN YAQRQVS... .........P SAPPMTEEMK
O_SN_99SE_  KQANFLGKYW PP.GGTRPGN YAQRQVS... .........P SAPPMTEEMK
U_CD___83C  RQANFLGKIW PSNKG.RPGN FLQNRPE... .........P TAPPAESFG.

            501                                                 550
00BW0762_1  .....FE... ..ETNPTP.. ...KQE.... PKDRE..... PLTSLKSLFG
00BW0768_2  .....FE... .ETTTPAP.. ...KQE.... LKDR.....E PLTALKSLFG
00BW0874_2  .....FE... ..ETTPAL.. ...KRE.... LKER.....E PLISLKSLFG
00BW1471_2  .....FE... ..ETTPAP.. ...KQE.... PKDR.....E PLTSLKSLFG
00BW1616_2  .....F.... .GETTPSP.. ...RQE.... AKDRE..... PLISLKSLFG
00BW1686_8  .....FE... ..ETTPAP.. ...KQE.... PKDR.....E PLTSLKSLFG
00BW1759_3  .....FE... ..ETTPAP.. ...KQE.... PKDRE..... TLTSLRSLFG
00BW1773_2  .....FE... ..ETTPAP.. ...KQE.... PKDRE..... PLTSLKSLFG
```

```
00BW1783_5  .....FE... ..ETTPVQ.. ...KQE.... TKDRE..... PLTSLKSLFG
00BW1795_6  .....F.... .EETTPSP.. ...KQE.... LKDKE..... PLTSLKSLFG
00BW1811_3  .....FE... ..ETTPAS.. ...KQE.... KKDRE..... TLTSLRSLFG
00BW1859_5  .....FE... ..ETTPAP.. ...KQE.... QKDR.....E PLTSLKSLFG
00BW1880_2  .....FE... ..ETTPAP.. ...KQE.... PKDRE..... PLTSLKSLFG
00BW1921_1  .....FE... ..ETTPAP.. ...KQE.... PKDR.....E PLTSLKSLFG
00BW2036_1  .....FE... ..ETTPAP.. ...KQE.... LKDR.....E PLISLKSLFG
00BW2063_6  .......... .EETTPAP.. ...KQE.... MKDKE..... PLISLKSLLG
00BW2087_2  .....FE... ..ETTPAS.. ...KQD.... LKDRE..... PLTSLKSLFG
00BW2127_2  .....FE... ..ETTHAP.. ...KQE.... LKDRE..... ALTSLKSLFG
00BW2128_3  .....FE... ..ETTPAP.. ...KQE.... PKNRE..... PLTSLKSLFG
00BW2276_7  .....FE... ..ETTPEL.. ...KQG.... PKDR.....E PLTSLKSLFG
00BW3819_3  .....FE... ..EITPAP.. ...KQE.... TKDRE..... PLTSLKSLFG
00BW3842_8  .....FE... ..ETTPAP.. ...KQE.... PKDRGPY.RE PLISLKSLFG
00BW3871_3  .....FE... ..ETTPVP.. ...KQE.... PTDR.....E PLTSLKSLFG
00BW3876_9  .....FE... ..ETTPTL.. ...KQE.... LKDRE..... PLTSLKSLFG
00BW3886_8  .....FE... ..ETTPVP.. ...KQE.... QKDRE..... ALTSLKSLFG
00BW3891_6  .....FE... ..EITPVP.. ...KQE.... PKDR.....E PLTSLKSLFG
00BW3970_2  .....FE... ..ETTPAP.. ...KQE.... PKDRE..... PLISLKSLFG
00BW5031_1  .....FG... ..ETTPAP.. ...KQE.... MKERE..... PLISLKSLFG
 96BW01B21  .....FE... ..ETTPAP.. ...KQE.... PKDRE..... PLTSLRSLFG
  96BW0407  .....FE... ..ETTPGQ.. ...KQE.... SKDRE..... TLTSLKSLFG
  96BW0502  .....FE... ..ETTPAP.. ...KQE.... PKDREPY.RE PLTALRSLFG
 96BW06_J4  .....FE... ..ETTPAL.. ...KQE.... PKDK.....E PLTSLKSPFG
 96BW11_06  .......... .EETTPAP.. ...KQE.... TKDRE..... PLISLKSLFG
  96BW1210  .....FE... ..ETTPAQ.. ...KQE.... PKDREP.... PLASLKSLFG
 96BW15B03  .....FE... ..ETTPAP.. ...KQE.... PKDR.....E PLISLKSLFG
 96BW16_26  .....FG... ..ETTPAP.. ...KQE.... PKDRE..... PLTSLRSLFG
 96BW17A09  .....FE... ..ETTPAP.. ...KQE.... PKDR.....E PLTSFKSLFG
 96BWMO1_5  .......... .EETTPAP.. ...KQE.... MKDKEPY.KE PLISLRSLFG
 96BWMO3_2  .....PE... .PTAPPAE.. ...RQE.... SKDR.....E PLISLKSLFG
98BWMC12_2  .....LE... ..ETTPAS.. ...KQE.... MKDR.....E PLISLKSLFG
98BWMC13_4  .......... .EETTPAP.. ...KQE.... PKDKE..... PLISLKSLFG
98BWMC14_a  .....FE... ..ETTPAP.. ...KQE.... QKDR.....E PLTSLKSLFG
98BWMO14_1  .....FESRP EPTAPPAES. ..FRQE.... PKDR.....E PLTALKSLFG
98BWMO18_d  .....FE... ..ETTPAL.. ...KQE.... PKDREA.... PLTSLKSLFG
98BWMO36_a  .....FE... ..ETNLAP.. ...KQE.... PKDRE..... PLTSLKSLFG
98BWMO37_d  .....FE... ..ETTPAP.. ...RQE.... AKDKE..... PLNSLKSLFG
99BW3932_1  .....FE... ..ETTPAP.. ...KQE.... LKDRE..... ALTSLKSLFG
99BW4642_4  .....FE... ..ETTPAP.. ...KQE.... PKDR.....E PLTSLKSLFG
99BW4745_8  .....FE... ..GATPTP.. ...KQE.... PRDR.....E PLTSLKSLFG
99BW4754_7  .....FE... ..ETTPTQ.. ...KQE.... SKDRE..... PLTSLKSLFG
99BWMC16_8  .....FE... ..ETNPAP.. ...KQE.... LKNRE..... TLTSLRSLFG
A2_CD_97CD  .......... ...EEITSSL ..KQE..... NREPS....T PAISLKSLFG
A2_CY_94CY  .......... .MGEEITSSL ..KQEL...E TREPY....N PAISLKSLFG
A2D___97KR  .......... .MGEETTPLQ ..KQEL...K NREQH....T PAISLKSLFG
A2G_CD_97C  .......... .......... .......... .......... ..........
A_BY_97BL0  .......... MGEEIT.... PSLK.Q...E QKDRE..QYP PSISLKSLFG
A_KE_Q23_A  .......... MGEETV.... SPLK.Q...E QKDRE..QAQ PLVSLKSLFG
A_SE_SE659  .......... .......... .......... .......... ..........
A_SE_SE725  .......... .......... .......... .......... ..........
A_SE_SE753  .......... MREEIA.... SPPK.Q...E Q..KG..QDP PLVSLKSLFG
A_SE_SE853  .......... .......... .......... .......... ..........
A_SE_SE889  .......... .......... .......... .......... ..........
A_SE_UGSE8  .......... MGEEIA.... SPPK.Q...E Q......NNP PSVSLKSLFG
A_UG_92UG0  .......... MREEIV.... SPPK.Q...E QNDRD..QNP PSVSLKSLFG
A_UG_U455_  .......... .MGEKMTSPA ..KQEL...K DREQ.....T PLVSLKSLFG
AC_IN_2130  .....FE... ..ETTPAL.. ...KQE.... QKDRE..... PLTSLKSLFG
AC_RW_92RW  ......ENFG MGEEIASPL. ..K.QE.... QKDRE..... PLISLKSLFG
AC_SE_SE94  .......... .......... .......... .......... ..........
ACD_SE_SE8  .......... FGEEITP... S.QK.Q...E QKDKE...LY PLASLKSLFG
```

```
ACG_BE_VI1  .......... KEDAIDSS.. ..PKQE.... PRDKG..LYP PLTSLKSLFG
AD_SE_SE69  .......... FGEEIAP... .SQKQE...Q KDK....ELY PLASLKSLFG
AD_SE_SE71  .......... .......... .......... .......... ..........
ADHK_NO_97  .......... IGEEIT.... SYQK.Q...E QKDRE..PPP PLVSLKSLFG
ADK_CD_MAL  .......... FGEEIK.... PSQK.Q...E QKDKE..L.Y PLASLKSLFG
AG_BE_VI11  .......... MEEEIT.... PSQK.Q...E PRDTG..LYP PLTSLKSLFG
AG_NG_92NG  .......... FGEEIAP... S.LK.Q...E PREKE..SPP L.TSLKSLFG
AGHU_GA_VI  .......... FGEEIA.... PSPR.P...E PREKE..R.Y PLTSLKSLFG
AGU_CD_Z32  .......... TKEEITS... S.PK.Q...E PRDKE..LYP PLASLKSLFG
AJ_BW_BW21  .......... FGEETA.... PSPK.Q...E GKDKE..L.Y PLTSLKSLFG
B_AU_VH_AF  .......... FGEETTTP.. .SQKQE.... PIDK...ELY PLASLRSLFG
B_CN_RL42_  .......... FGEETTTP.. .SQKQE.... PIDK...ELY PLASLKSLFG
B_DE_D31_U  .......... FGEETATP.. .FQKQE.... PIDK...ELY PLASLRSLFG
B_DE_HAN_U  .......... FGEATAP... .SQKQE.... PIDK...ELY PLASLKSLFG
B_FR_HXB2_  .......... SGVETTTP.. .PQKQE.... PIDK...ELY PLTSLRSLFG
B_GA_OYI__  .......... FGEETTTP.. .PQKQE.... PIDK...GLY PLTSLRSLFG
B_GB_CAM1_  .......... FGEEKTTP.. .SQKQE.... PIDK...ELY PLASLRSLFG
B_GB_GB8_A  .......... FGGETTTP.. .SQKQE.... PINK...EPY PLASLRSLFG
B_GB_MANC_  .......... FGEETTTP.. .AQKQE.... PIDK...ELY PLASLRPLFG
B_KR_WK_AF  .......... FGEETTTP.. .SQKQE.... PIDK...ELY PLASLRSLFG
B_NL_3202A  .......... FGEETTTP.. .SQKQE.... PRDK...ELY PLASLRSLFG
B_TW_TWCYS  .......... FGEQTTTP.. .SQKQE.... PIDK...DLY PLASLESLFG
B_US_BC_L0  .......... FGEETTTP.. .PQKQERE.. ..DK...EMY PLASLRSLFG
B_US_DH123  .......... FGEETATP.. .SQKQE.... ..PK...ELY PLASLKSLFG
B_US_JRCSF  .......... FGEETATP.. .SQKQEQKQE PIDK...ELY PLTSLRSLFG
B_US_MNCG_  .......... FGEETTTP.. .YQKQEKKQE TIDK...DLY PLASLKSLFG
B_US_P896_  .......... FGEETTTP.. .SQKQE.... PIDK...ELY PLASLRSLFG
B_US_RF_M1  .......... FGEETTP... .SQKQE.... KIDK...ELY PLASLKSLFG
B_US_SF2_K  .......... FGEEKTTP.. .SQKQE.... PIDK...ELY PLTSLRSLFG
B_US_WEAU1  .......... FREETTTP.. .SQKQE.... PIDK...ELY PLTSLKSLFG
B_US_WR27_  .......... FGXETTTP.. .SQKQE.... PIDK...ELY PLASLRSLFV
B_US_YU2_M  .......... FGEETTTP.. .SQKQE.... PIDK...ELY PLASLRSLFG
BF1_BR_93B  .......... FGEEVTTP.. .SQKQE.... PIDK...EMY PLASLRSLFG
C_BR_92BR0  ......ESFR FGEETTTPS. ..RKQE.... TIDKEL.... PLTSLKSLFG
C_BW_96BW0  .....FE... ..ETTPVP.. ...KQE.... PKDRE..... PLTSLKSLFG
C_BW_96BW1  .......... .EETTPAP.. ...KQE.... TKDRE..... PLISLKSLFG
C_BW_96BW1  .....FE... ..ETTPAQ.. ...KQE.... PKDREP.... PLASLKSLFG
C_BW_96BW1  .....FE... ..ETTPAP.. ...KQE.... PKDR.....E PLISLKSLFG
C_ET_ETH22  PTAPPPESFR FEEATPSPK. ..Q..E.... LKDRE..... ALTSLKSLFG
C_IN_93IN1  .....FE... ..ETTPAP.. ...KQE.... PKDRE..... PLTSLKSLFG
C_IN_93IN9  .....FE... ..ETPPAP.. ...KQE.... PKDRE..... PLTSLRSLFG
C_IN_93IN9  .....FE... ..ETTPAL.. ...KQE.... PKDRE..... PLTSLKSLFG
C_IN_94IN1  .....FE... ..ETPPAP.. ...KQE.... PKERE..... PLTSLRSLFG
C_IN_95IN2  .....FE... ..ETTPAP.. ...KQE.... PKDRE..... PLTSLRSLFG
CRF01_AE_C  .......... MGEEIT.... SFPK.Q...E QKDKE..HPS PLVSLKSLFG
CRF01_AE_C  .......... MGEEIT.... SFPK.Q...E QKDKK..QPP PLVSLKSLFG
CRF01_AE_C  .......... MGEEIT.... SFSR.Q...E QKDRE..HPP PLVSLKSLFG
CRF01_AE_T  .......... MGEETT.... .SLLKQ...E QKDKE..HHP PLVSLKSLFG
CRF01_AE_T  .......... MGEEIT.... SLPK.Q...E QKDKD..PPP .LVSLKSLFG
CRF01_AE_T  .......... MGEE...... .......... QKDKE..HPP PSVSLKSLFG
CRF01_AE_T  .......... MGEETT.... SSLK.Q...E QKDKE..PPP PLISLKSLFG
CRF01_AE_T  .......... MGEEITGEEI TSLPKQ...E QKDKE..HPP PLVSLKSLFG
CRF01_AE_T  .......... MGEEIT.... SFLK.Q...E QKDKE..HPP PSVSLKSLFG
CRF02_AG_F  .......... MGEEIT.... SPPK.Q...E ARDQG..LYP PLASLKSLFG
CRF02_AG_F  .......... MGEEIT.... SPPK.Q...E PRDQG..LYP PLASLKSLFG
CRF02_AG_G  .......... TREEITSS.. ..PQQE.... PRDKG..LYP PLTSLKSLFG
CRF02_AG_N  .......... MGEEIP.... PSPQ.Q...E PRDKG..LYP PLTSLKSLFG
CRF02_AG_S  .......... IGEEIT.... SSQK.Q...E PGDKG..LYP PLASLKSLFG
CRF02_AG_S  .......... MGEEIT.... SSPK.Q...E PGDKG..LYP PLTSLKSLFG
CRF03_AB_R  .......... MGEEIT.... PSLK.Q...E QKDRE..QHP PSISLKSLFG
CRF03_AB_R  .......... MGEEIT.... PSLK.Q...E QKDRG..QHP PSISLKSLFG
```

```
CRF04_cpx_  .......... RKEETTS... S.LK.Q...E PRDKE..LYP .LTSLKSLFG
CRF04_cpx_  .......... MKEETTS... S.PK.Q...E PRDKE..LYP .LTSLKSLFG
CRF04_cpx_  .......... MKEETTS... S.PK.Q...E QRDKE..LYP .ITSLKSLFG
CRF05_DF_B  .......... FGEEIAS... .SPKQE...Q KDEG...LYP PLASLKSLFG
CRF05_DF_B  .......... FGEEITP... .SPKQE...Q KDEG...KYP PLASLKSLFG
CRF06_cpx_  .......... FGEEIAP... S.PK.Q...E SKEKEEKGLY PLASLKSLFG
CRF06_cpx_  .......... FGEETAP... S.PE.Q...K PKEKE...LY PLTSLRSLFG
CRF06_cpx_  .......... FGEETAP... S.LK.Q...E PKEKEKE.LY PLASLKSLFG
CRF06_cpx_  .......... FGEEIAP... S.PK.Q...E PKEKE...LY PLASLKSLFG
CRF11_cpx_  .......... FGEEIAP... .SPK.Q...E PKEKEK.ELY PLTSLKSLFG
CRF11_cpx_  .......... FGEETTP... .SPK.Q...E PKEK...ELY PITSLKSLFG
D_CD_84ZR0  .......... FGEEITP... .SQKQEQK.. DKDK...ELY PLASLKSLFG
D_CD_ELI_K  .......... FGEEITP... .SQKQE...Q KDK....ELY PLTSLKSLFG
D_CD_NDK_M  .......... FGEEITP... .SQKQE...Q KDK....ELY PLASLKSLFG
D_UG_94UG1  .......... LGEEITP... .PQKQE...Q KDK....ELY PLTSLKSLFG
F1_BE_VI85  .....FR... .EEITPSP.. ...KQE.... QKDGEL..YP PLASLKSLFG
F1_BR_93BR  .....FG... .EETTPSP.. ...KQE.... QKDEGL..YP PLASLKSLFG
F1_FI_FIN9  .....IR... .EEVTPSP.. ...RQE.... QKEEGQ..YP PLASLKSLFG
F1_FR_MP41  .....FK... .EEITPSP.. ...KQE.... QKDEGQGLYP PLASLKSLFG
F2_CM_MP25  .....FG... .EEIAPSP.. ...KQE.... QKDKEQ..VP PLISLKSLFG
F2KU_BE_VI  .....FG... .EEINPSP.. ...RQE.... TKDKGQ..EP PLTSLKSLFG
G_BE_DRCBL  .......... FGEEIAP... S.PK.Q...E QKEKE..LYP L.SSLKSLFG
G_NG_92NG0  .......... FGEEIAP... S.PK.Q...E PKEKE..LYP L.TSLKSLFG
G_SE_SE616  .......... FGEEIAP... S.PK.Q...E MKEKE..LYP ...SLKSLFG
H_BE_VI991  .....FG... .EEITPSP.. ...RQE.... LKEQE....P PLTSLRSLFG
H_BE_VI997  .....FG... .EEMTSSP.. ...KQE.... LKDKE....P PFASLKSLFG
H_CF_90CF0  .....FG... .EEMTPSP.. ...KQEQ... LKDKE....P PLASLRSLFG
J_SE_SE702  .....FG... ..EEIPSP.. ...KQE.... PKDKE...LY PLTSLRSLFG
J_SE_SE788  .....LG... ..EEIPSP.. ...KQE.... PKDKE...LY PLTSLKSLFG
K_CD_EQTB1  .....FG... .EKITPSL.. ...RQE.... MKDQEQ..GP PLTSLKSLFG
K_CM_MP535  .....FG... .EEITPSP.. ...RQE.... TKDKEQ..SP PLTSLKSLFG
N_CM_YBF30  .....FQ... ..EEKSTQ.. GKEMQE...N QERTENSLYP PLTSLRSLFG
O_CM_ANT70  .......... .GQENQEQ.. ...KGG.... PNE.....LY PFASLKSLFG
O_CM_MVP51  .......... .EQENQSQ.. ...KGD.... QEE.....LY PFASLKSLFG
O_SN_99SE_  .......... .GQENQEQ.. ...KED.... QNE.....LY PFASLRSLFG
O_SN_99SE_  .......... .GQENQEQ.. ...KGD.... QNE.....LY PFASLKSLFG
U_CD___83C  .....FG... .EETTPSP.. ...KQE.... PRDKESL.YP PLTSLKSLFG

            551
00BW0762_1  SDPLSQ
00BW0768_2  SDPLSQ
00BW0874_2  NDPLSQ
00BW1471_2  SDPLSQ
00BW1616_2  SDPLSQ
00BW1686_8  SDPLSQ
00BW1759_3  SDPLSQ
00BW1773_2  SDPLSQ
00BW1783_5  SDPLSQ
00BW1795_6  SDPLSQ
00BW1811_3  SDPLSQ
00BW1859_5  SDPLSQ
00BW1880_2  NDPLSQ
00BW1921_1  SDPLSQ
00BW2036_1  SDPLSQ
00BW2063_6  NDPLSQ
00BW2087_2  SDPLSQ
00BW2127_2  SDPLSQ
00BW2128_3  SDPWSQ
00BW2276_7  SDPLSQ
00BW3819_3  SDPLSQ
00BW3842_8  SDPLSQ
```

```
00BW3871_3  SDPLSQ
00BW3876_9  SDPLSQ
00BW3886_8  SDPLSQ
00BW3891_6  SDPLSQ
00BW3970_2  SDPLSQ
00BW5031_1  SDPLSQ
 96BW01B21  SDPLSQ
  96BW0407  NDPLSQ
  96BW0502  SGPLSQ
 96BW06_J4  SDPLSQ
 96BW11_06  SDPLSQ
  96BW1210  NDPLSQ
 96BW15B03  SDPLSQ
 96BW16_26  NDPLSQ
 96BW17A09  SDPLSQ
 96BWMO1_5  SDPLSQ
 96BWMO3_2  SDPLSQ
98BWMC12_2  NDPLSQ
98BWMC13_4  SDPLSQ
98BWMC14_a  NDPLSQ
98BWMO14_1  SDPLSQ
98BWMO18_d  SDPLSQ
98BWMO36_a  SDPLSQ
98BWMO37_d  SDPLSQ
99BW3932_1  SDPLSQ
99BW4642_4  SDPLSQ
99BW4745_8  SDPLSQ
99BW4754_7  NDPLSQ
99BWMC16_8  GDPLSQ
A2_CD_97CD  NDLLSQ
A2_CY_94CY  NDPLLQ
A2D___97KR  NDPLLQ
A2G_CD_97C  ......
A_BY_97BL0  NDPLSQ
A_KE_Q23_A  NDLLSQ
A_SE_SE659  ......
A_SE_SE725  ......
A_SE_SE753  NDLLSQ
A_SE_SE853  ......
A_SE_SE889  ......
A_SE_UGSE8  NDLLSQ
A_UG_92UG0  NDLLSQ
A_UG_U455_  NDPLSQ
AC_IN_2130  SDPLSQ
AC_RW_92RW  NDPLSQ
AC_SE_SE94  ......
ACD_SE_SE8  NDP...
ACG_BE_VI1  NDP...
AD_SE_SE69  NDP...
AD_SE_SE71  ......
ADHK_NO_97  NDPLSQ
ADK_CD_MAL  NDQLSQ
AG_BE_VI11  NDP...
AG_NG_92NG  NDP...
AGHU_GA_VI  SDP...
AGU_CD_Z32  SDP...
AJ_BW_BW21  SDP...
B_AU_VH_AF  NDPSSQ
B_CN_RL42_  NDPSSQ
B_DE_D31_U  NDPSSQ
B_DE_HAN_U  SDPSSQ
```

B_FR_HXB2_	NDPSSQ
B_GA_OYI__	NDPSSQ
B_GB_CAM1_	NDPSSQ
B_GB_GB8_A	NDPSSQ
B_GB_MANC_	NDPSSQ
B_KR_WK_AF	NDPSSQ
B_NL_3202A	NDPSSQ
B_TW_TWCYS	NDPSSQ
B_US_BC_L0	NDPSSQ
B_US_DH123	NDP...
B_US_JRCSF	NDPSSQ
B_US_MNCG_	NDPLSQ
B_US_P896_	NDPSSQ
B_US_RF_M1	NDPSSQ
B_US_SF2_K	NDPSSQ
B_US_WEAU1	NDPSSQ
B_US_WR27_	NDPSSQ
B_US_YU2_M	SDPSSQ
BF1_BR_93B	NDPSSQ
C_BR_92BR0	SDPLST
C_BW_96BW0	SDPLSQ
C_BW_96BW1	SDPLSQ
C_BW_96BW1	NDPLSQ
C_BW_96BW1	SDPLSQ
C_ET_ETH22	NDHLLQ
C_IN_93IN1	SDLLSQ
C_IN_93IN9	SDPLSQ
C_IN_93IN9	SDPLSQ
C_IN_94IN1	SDPLSQ
C_IN_95IN2	SDPLSQ
CRF01_AE_C	NDPLSQ
CRF01_AE_C	NDPLSQ
CRF01_AE_C	NDPLSQ
CRF01_AE_T	NDPSSQ
CRF01_AE_T	NDPLSQ
CRF01_AE_T	NDPLSQ
CRF01_AE_T	NDPLSQ
CRF01_AE_T	NDPLSQ
CRF01_AE_T	NDPLSQ
CRF02_AG_F	NDP...
CRF02_AG_F	NDP...
CRF02_AG_G	NDP...
CRF02_AG_N	NDP...
CRF02_AG_S	NDP...
CRF02_AG_S	NDPYSQ
CRF03_AB_R	DDPLSQ
CRF03_AB_R	NDPLSQ
CRF04_cpx_	SDPLSQ
CRF04_cpx_	NHPLSQ
CRF04_cpx_	SDPLSR
CRF05_DF_B	NDPLSQ
CRF05_DF_B	NDPLSQ
CRF06_cpx_	SDP...
CRF06_cpx_	NDP...
CRF06_cpx_	NDP...
CRF06_cpx_	SDP...
CRF11_cpx_	SDP...
CRF11_cpx_	SDPLSQ
D_CD_84ZR0	NDPLSQ
D_CD_ELI_K	NDPLSQ
D_CD_NDK_M	NDPSSQ

```
D_UG_94UG1   NDPLSQ
F1_BE_VI85   NDP...
F1_BR_93BR   NDP...
F1_FI_FIN9   NDP...
F1_FR_MP41   SDP...
F2_CM_MP25   SDQ...
F2KU_BE_VI   SDPLLQ
G_BE_DRCBL   NDQ...
G_NG_92NG0   SDP...
G_SE_SE616   SDP...
H_BE_VI991   NDQ...
H_BE_VI997   NDPLSQ
H_CF_90CF0   SDPLLQ
J_SE_SE702   SDPLSQ
J_SE_SE788   SDPLSQ
K_CD_EQTB1   SDPLSQ
K_CM_MP535   NDPLSQ
N_CM_YBF30   NDPSSQ
O_CM_ANT70   TDQ...
O_CM_MVP51   TDQ...
O_SN_99SE_   TDQ...
O_SN_99SE_   TDQ...
U_CD___83C   SDPSLQ
```

Table 12. HIV Env Sequence Alignment
GCG Multiple Sequence File.
Written by Omiga 1.1

Name: 00BW0762_1	Len: 962	Check: 4645	Weight:	1.00
Name: 00BW0768_2	Len: 962	Check: 9565	Weight:	1.00
Name: 00BW0874_2	Len: 962	Check: 7745	Weight:	1.00
Name: 00BW1471_2	Len: 962	Check: 9593	Weight:	1.00
Name: 00BW1616_2	Len: 962	Check: 792	Weight:	1.00
Name: 00BW1686_8	Len: 962	Check: 3744	Weight:	1.00
Name: 00BW1759_3	Len: 962	Check: 9808	Weight:	1.00
Name: 00BW1773_2	Len: 962	Check: 3500	Weight:	1.00
Name: 00BW1783_5	Len: 962	Check: 9684	Weight:	1.00
Name: 00BW1795_6	Len: 962	Check: 8410	Weight:	1.00
Name: 00BW1811_3	Len: 962	Check: 2068	Weight:	1.00
Name: 00BW1859_5	Len: 962	Check: 5692	Weight:	1.00
Name: 00BW1880_2	Len: 962	Check: 1901	Weight:	1.00
Name: 00BW1921_1	Len: 962	Check: 5923	Weight:	1.00
Name: 00BW2036_1	Len: 962	Check: 7035	Weight:	1.00
Name: 00BW2063_6	Len: 962	Check: 4853	Weight:	1.00
Name: 00BW2087_2	Len: 962	Check: 2085	Weight:	1.00
Name: 00BW2127_2	Len: 962	Check: 4015	Weight:	1.00
Name: 00BW2128_3	Len: 962	Check: 5884	Weight:	1.00
Name: 00BW2276_7	Len: 962	Check: 8913	Weight:	1.00
Name: 00BW3819_3	Len: 962	Check: 9390	Weight:	1.00
Name: 00BW3842_8	Len: 962	Check: 8867	Weight:	1.00
Name: 00BW3871_3	Len: 962	Check: 7069	Weight:	1.00
Name: 00BW3876_9	Len: 962	Check: 4761	Weight:	1.00
Name: 00BW3886_8	Len: 962	Check: 7681	Weight:	1.00
Name: 00BW3891_6	Len: 962	Check: 379	Weight:	1.00
Name: 00BW3970_2	Len: 962	Check: 8001	Weight:	1.00
Name: 00BW5031_1	Len: 962	Check: 4902	Weight:	1.00
Name: 96BW01B21	Len: 962	Check: 5774	Weight:	1.00
Name: 96BW0407	Len: 962	Check: 4260	Weight:	1.00
Name: 96BW0502	Len: 962	Check: 4658	Weight:	1.00
Name: 96BW06_J4	Len: 962	Check: 9749	Weight:	1.00
Name: 96BW11_06	Len: 962	Check: 4328	Weight:	1.00
Name: 96BW1210	Len: 962	Check: 3855	Weight:	1.00
Name: 96BW15B03	Len: 962	Check: 9133	Weight:	1.00
Name: 96BW16_26	Len: 962	Check: 5	Weight:	1.00
Name: 96BW17A09	Len: 962	Check: 6458	Weight:	1.00
Name: 96BWMO1_5	Len: 962	Check: 9487	Weight:	1.00
Name: 96BWMO3_2	Len: 962	Check: 8766	Weight:	1.00
Name: 98BWMC12_2	Len: 962	Check: 2722	Weight:	1.00
Name: 98BWMC13_4	Len: 962	Check: 2526	Weight:	1.00
Name: 98BWMC14_a	Len: 962	Check: 7761	Weight:	1.00
Name: 98BWMO14_1	Len: 962	Check: 93	Weight:	1.00
Name: 98BWMO18_d	Len: 962	Check: 279	Weight:	1.00
Name: 98BWMO36_a	Len: 962	Check: 134	Weight:	1.00
Name: 98BWMO37_d	Len: 962	Check: 9669	Weight:	1.00
Name: 99BW3932_1	Len: 962	Check: 3527	Weight:	1.00
Name: 99BW4642_4	Len: 962	Check: 1175	Weight:	1.00
Name: 99BW4745_8	Len: 962	Check: 8117	Weight:	1.00
Name: 99BW4754_7	Len: 962	Check: 5709	Weight:	1.00
Name: 99BWMC16_8	Len: 962	Check: 285	Weight:	1.00
Name: A2_CD_97CD	Len: 962	Check: 2892	Weight:	1.00
Name: A2_CY_94CY	Len: 962	Check: 8628	Weight:	1.00
Name: A2D___97KR	Len: 962	Check: 471	Weight:	1.00
Name: A2G_CD_97C	Len: 962	Check: 939	Weight:	1.00
Name: A_BY_97BL0	Len: 962	Check: 4291	Weight:	1.00
Name: A_KE_Q23_A	Len: 962	Check: 1190	Weight:	1.00

Name:	A_SE_SE659	Len:	962	Check:	6674	Weight:	1.00
Name:	A_SE_SE725	Len:	962	Check:	4925	Weight:	1.00
Name:	A_SE_SE753	Len:	962	Check:	2482	Weight:	1.00
Name:	A_SE_SE853	Len:	962	Check:	1860	Weight:	1.00
Name:	A_SE_SE889	Len:	962	Check:	2102	Weight:	1.00
Name:	A_SE_UGSE8	Len:	962	Check:	5063	Weight:	1.00
Name:	A_UG_92UG0	Len:	962	Check:	6685	Weight:	1.00
Name:	A_UG_U455_	Len:	962	Check:	8657	Weight:	1.00
Name:	AC_IN_2130	Len:	962	Check:	7784	Weight:	1.00
Name:	AC_RW_92RW	Len:	962	Check:	4676	Weight:	1.00
Name:	AC_SE_SE94	Len:	962	Check:	2949	Weight:	1.00
Name:	ACD_SE_SE8	Len:	962	Check:	1464	Weight:	1.00
Name:	ACG_BE_VI1	Len:	962	Check:	2980	Weight:	1.00
Name:	AD_SE_SE69	Len:	962	Check:	8959	Weight:	1.00
Name:	AD_SE_SE71	Len:	962	Check:	7056	Weight:	1.00
Name:	ADHK_NO_97	Len:	962	Check:	487	Weight:	1.00
Name:	ADK_CD_MAL	Len:	962	Check:	2555	Weight:	1.00
Name:	AG_BE_VI11	Len:	962	Check:	6342	Weight:	1.00
Name:	AG_NG_92NG	Len:	962	Check:	1272	Weight:	1.00
Name:	AGHU_GA_VI	Len:	962	Check:	1974	Weight:	1.00
Name:	AGU_CD_Z32	Len:	962	Check:	4356	Weight:	1.00
Name:	AJ_BW_BW21	Len:	962	Check:	9995	Weight:	1.00
Name:	B_AU_VH_AF	Len:	962	Check:	5833	Weight:	1.00
Name:	B_CN_RL42_	Len:	962	Check:	4092	Weight:	1.00
Name:	B_DE_D31_U	Len:	962	Check:	5486	Weight:	1.00
Name:	B_DE_HAN_U	Len:	962	Check:	3480	Weight:	1.00
Name:	B_FR_HXB2_	Len:	962	Check:	6939	Weight:	1.00
Name:	B_GA_OYI__	Len:	962	Check:	9780	Weight:	1.00
Name:	B_GB_CAM1_	Len:	962	Check:	9716	Weight:	1.00
Name:	B_GB_GB8_C	Len:	962	Check:	4180	Weight:	1.00
Name:	B_GB_MANC_	Len:	962	Check:	9762	Weight:	1.00
Name:	B_KR_WK_AF	Len:	962	Check:	6641	Weight:	1.00
Name:	B_NL_3202A	Len:	962	Check:	7168	Weight:	1.00
Name:	B_TW_TWCYS	Len:	962	Check:	3591	Weight:	1.00
Name:	B_US_BC_L0	Len:	962	Check:	7266	Weight:	1.00
Name:	B_US_DH123	Len:	962	Check:	6905	Weight:	1.00
Name:	B_US_JRCSF	Len:	962	Check:	9381	Weight:	1.00
Name:	B_US_MNCG_	Len:	962	Check:	9951	Weight:	1.00
Name:	B_US_P896_	Len:	962	Check:	5855	Weight:	1.00
Name:	B_US_RF_M1	Len:	962	Check:	6075	Weight:	1.00
Name:	B_US_SF2_K	Len:	962	Check:	1434	Weight:	1.00
Name:	B_US_WEAU1	Len:	962	Check:	5451	Weight:	1.00
Name:	B_US_WR27_	Len:	962	Check:	4262	Weight:	1.00
Name:	B_US_YU2_M	Len:	962	Check:	5841	Weight:	1.00
Name:	BF1_BR_93B	Len:	962	Check:	5506	Weight:	1.00
Name:	C_BR_92BR0	Len:	962	Check:	8769	Weight:	1.00
Name:	C_BW_96BW0	Len:	962	Check:	6197	Weight:	1.00
Name:	C_BW_96BW1	Len:	962	Check:	8144	Weight:	1.00
Name:	C_BW_96BW1	Len:	962	Check:	1160	Weight:	1.00
Name:	C_BW_96BW1	Len:	962	Check:	2736	Weight:	1.00
Name:	C_ET_ETH22	Len:	962	Check:	8219	Weight:	1.00
Name:	C_IN_93IN1	Len:	962	Check:	4068	Weight:	1.00
Name:	C_IN_93IN9	Len:	962	Check:	3674	Weight:	1.00
Name:	C_IN_93IN9	Len:	962	Check:	1581	Weight:	1.00
Name:	C_IN_94IN1	Len:	962	Check:	9352	Weight:	1.00
Name:	C_IN_95IN2	Len:	962	Check:	6988	Weight:	1.00
Name:	CRF01_AE_C	Len:	962	Check:	8684	Weight:	1.00
Name:	CRF01_AE_C	Len:	962	Check:	3342	Weight:	1.00
Name:	CRF01_AE_C	Len:	962	Check:	5017	Weight:	1.00
Name:	CRF01_AE_T	Len:	962	Check:	9124	Weight:	1.00
Name:	CRF01_AE_T	Len:	962	Check:	2718	Weight:	1.00

```
Name: CRF01_AE_T      Len:   962  Check: 2104  Weight:   1.00
Name: CRF01_AE_T      Len:   962  Check: 8495  Weight:   1.00
Name: CRF01_AE_T      Len:   962  Check: 4076  Weight:   1.00
Name: CRF01_AE_T      Len:   962  Check: 948  Weight:   1.00
Name: CRF02_AG_F      Len:   962  Check: 9298  Weight:   1.00
Name: CRF02_AG_F      Len:   962  Check: 9278  Weight:   1.00
Name: CRF02_AG_G      Len:   962  Check: 4373  Weight:   1.00
Name: CRF02_AG_N      Len:   962  Check: 8955  Weight:   1.00
Name: CRF02_AG_S      Len:   962  Check: 252  Weight:   1.00
Name: CRF02_AG_S      Len:   962  Check: 5147  Weight:   1.00
Name: CRF03_AB_R      Len:   962  Check: 2239  Weight:   1.00
Name: CRF03_AB_R      Len:   962  Check: 2671  Weight:   1.00
Name: CRF04_cpx_      Len:   962  Check: 4892  Weight:   1.00
Name: CRF04_cpx_      Len:   962  Check: 8070  Weight:   1.00
Name: CRF04_cpx_      Len:   962  Check: 5453  Weight:   1.00
Name: CRF05_DF_B      Len:   962  Check: 174  Weight:   1.00
Name: CRF05_DF_B      Len:   962  Check: 2694  Weight:   1.00
Name: CRF06_cpx_      Len:   962  Check: 7351  Weight:   1.00
Name: CRF06_cpx_      Len:   962  Check: 5073  Weight:   1.00
Name: CRF06_cpx_      Len:   962  Check: 661  Weight:   1.00
Name: CRF06_cpx_      Len:   962  Check: 8440  Weight:   1.00
Name: CRF11_cpx_      Len:   962  Check: 2217  Weight:   1.00
Name: CRF11_cpx_      Len:   962  Check: 8216  Weight:   1.00
Name: D_CD_84ZR0      Len:   962  Check: 4843  Weight:   1.00
Name: D_CD_ELI_K      Len:   962  Check: 8403  Weight:   1.00
Name: D_CD_NDK_M      Len:   962  Check: 5813  Weight:   1.00
Name: D_UG_94UG1      Len:   962  Check: 9407  Weight:   1.00
Name: F1_BE_VI85      Len:   962  Check: 2982  Weight:   1.00
Name: F1_BR_93BR      Len:   962  Check: 8919  Weight:   1.00
Name: F1_FI_FIN9      Len:   962  Check: 6761  Weight:   1.00
Name: F1_FR_MP41      Len:   962  Check: 478  Weight:   1.00
Name: F2_CM_MP25      Len:   962  Check: 9292  Weight:   1.00
Name: F2KU_BE_VI      Len:   962  Check: 567  Weight:   1.00
Name: G_BE_DRCBL      Len:   962  Check: 6261  Weight:   1.00
Name: G_NG_92NG0      Len:   962  Check: 4508  Weight:   1.00
Name: G_SE_SE616      Len:   962  Check: 6733  Weight:   1.00
Name: H_BE_VI991      Len:   962  Check: 7498  Weight:   1.00
Name: H_BE_VI997      Len:   962  Check: 8345  Weight:   1.00
Name: H_CF_90CF0      Len:   962  Check: 2490  Weight:   1.00
Name: J_SE_SE702      Len:   962  Check: 4446  Weight:   1.00
Name: J_SE_SE788      Len:   962  Check: 1662  Weight:   1.00
Name: K_CD_EQTB1      Len:   962  Check: 7406  Weight:   1.00
Name: K_CM_MP535      Len:   962  Check: 512  Weight:   1.00
Name: N_CM_YBF30      Len:   962  Check: 1733  Weight:   1.00
Name: O_CM_ANT70      Len:   962  Check: 75  Weight:   1.00
Name: O_CM_MVP51      Len:   962  Check: 3290  Weight:   1.00
Name: O_SN_99SE_      Len:   962  Check: 6963  Weight:   1.00
Name: O_SN_99SE_      Len:   962  Check: 6278  Weight:   1.00
Name: U_CD___83C      Len:   962  Check: 9044  Weight:   1.00

                1                                                  50
00BW0762_1  ....MRVMGI MRNC.QQWWI WV.ILGFWML MVCN.VIGNL WVTVYYGVPV
00BW0768_2  ....MRVREI LRNC.QQWWT WG.SLGFWMV MIYS.VVGEL WVTVYYGVPV
00BW0874_2  ....MRAMGT QRNC.RQWWI WG.ILGFWML MTCS.GVG.E MVTVYYGVPV
00BW1471_2  ....MRVMGI LRSC.QQWWI WG.ILGFWML MICS.VLGNL WVTVYYGVPV
00BW1616_2  ....MRVMGI QRNC.QRWWI WG.ILGFWMI Y..N.VVGNL WVTVYYGVPV
00BW1686_8  ....MRVKGI QRNW.PQWWI WG.SLGFWML MFYS.VMGNL WVTVYYGVPV
00BW1759_3  ....MRVRGI PRNW.QQWWI WG.ILGFCMI ITCK.VVGNL WVTVYYGVPV
00BW1773_2  ....MRVREI LRSY.QHWWM WS.ILGLWIL IISN.VVGNL WVTVYYGVPV
00BW1783_5  ....MRVMGI KRNC.PPWWI WG.ILGFWML MICN.VMGNL WVTVYYGVPV
00BW1795_6  ....TRVMGI RRNW.QQWWI WG.ILGFWML IICN.VMGNM WVTVYYGVPV
```

```
00BW1811_3  ....MRVRGT LKNY.QRWWI WG.ILGLWIL LINI.VVGNL WVTVYYGVPV
00BW1859_5  ....MRVRGI PRNW.QQWWI WG.ILGFWIL MICN.VVGNL WVTVYYGVPV
00BW1880_2  ....MRVTGI MRNC.QQWWI WV.ILGFWML MICN.VIGNL WVTVYYGVPV
00BW1921_1  ....MRVRGI QRNW.QQWWI WG.SLGFWVM A.CS.VVGNL WVTVYYGVPV
00BW2036_1  ....MRVRGI SRNW.QQWWI WG.ILGFWMF MICS.VLGNL WVTVYYGVPV
00BW2063_6  ....MRVMGI MRNW.PPWWI WG.ILGFWML MICN.VMGNL WVTVYYGVPV
00BW2087_2  ....MRVTGM WKNC.QQWWI WG.ILGFWML MICS.VLGNL WVTVYYGVPV
00BW2127_2  ....MRVRGI PRNW.QQWWI WG.ILGFW.. MIYS.MMGNL WVTVYYGVPV
00BW2128_3  ....MRVRGI LKNC.QQWWI WI.ILGFWLL IITN.VVGKL WVTVYYGVPV
00BW2276_7  ....MRVRGI LRNW.QQWWI WG.ILGFWMV MICS.VWGNL WVTVYYGVPV
00BW3819_3  ....MRVRGI LRNW.QQWWI WG.ILSFWVL MICS.RGEDR WVTVYYGVPV
00BW3842_8  ....MRVRGI LRNW.QQWWI WV.ILGFW.. .IYS.VAGNL WVTVYYGVPV
00BW3871_3  ....MRVRGI QRNW.QQWWI WG.SLGFWML MIYN.VMGSL WVTVYYGVPV
00BW3876_9  ....MRVREI LRNW.KQLWT WG.ILGFWVL IICS.AGGNL WVTVYYGVPV
00BW3886_8  ....MRVRGI LRIW.QWWWI WA.SLGFWML IICN.EKGKL WVTVYYGVPV
00BW3891_6  ....MRVRGI LRNY.QQWWI WG.ILGFWML MMCN.VMGDL WVTVYYGVPV
00BW3970_2  ....MRVKGI MRNC.QQWWI WG.ILGFWML LICN.GEGNL WVTVYYGVPV
00BW5031_1  ....MRVMGT QRNC.QQWWI WG.ILGFWML MIYN.VGGNL WVTVYYGVPV
 96BW01B21  ....MRVRGI LRNY.PQWWI WG.ILGFWMI MLCN.VMGNL WVTVYYGVPV
  96BW0407  ....MRVMGI QRNC.QQWWI WG.ILGFWMI FNGM.GSW.. VT.VYYGVPV
  96BW0502  ....MRVMGI LKNY.QQWWM WG.ILGFWML IISS.VVGNL WVTVYYGVPV
 96BW06_J4  ....MRVKGI PRNW.QQWWI WG.SLGFWII C..S.VMGNL WVTVYYGVPV
 96BW11_06  ....MRVMEI MRNC.QQWWI WG.ILGFWML MICN.VMGKS WVTVYYGVPV
  96BW1210  ....MRVRGI LRNY.LQWWI WG.ILGFWML MVCS.K.ENM WVTVYG.VPV
 96BW15B03  ....MRVRGI LRSW.QQWWI WG.TLGFW.. .ICS.GLGNL WVTVYDGVPV
 96BW16_26  ....MIVRGI LKTC.QQWWI WI.ILGFWIL IINN.VVGHL WVTVYDGVPV
 96BW17A09  ....MRVMGI LRNC.QQWWI WG.ILGFWML MICS.VLGNL WVTVYYGVPV
 96BWMO1_5  ....MRVMGI KKNW.QPWWI WG.VLGFWTL MICS.VMGNL WVTVYYGVPV
 96BWMO3_2  ....MRVRGT QRNW.QRWWI WS.ILAFWIL INCN.GEEKL WVTVYYGVPV
98BWMC12_2  ....MRVMGI QKNC.QRWWI WG.ILGFWMI MSYS.VLGNL WVTVYYGVPV
98BWMC13_4  ....MRVMGI KMNW.QQWWI WG.ILGFWML MICS.VMGNL WVTVYYGVPV
98BWMC14_a  ....MRVKGI LRNW.LQWWI WG.SLGFWML C..S.VMGNM WVTVYYGVPV
98BWMO14_1  ....MRVMGT LRNC.QQWWT WG.ILGFWML MICS.VGGNL WVTVYYGVPV
98BWMO18_d  ....MRVMGI QKNC.QHWWI WG.ILGFWML MICN.GK.DL WVTVYYGVPV
98BWMO36_a  ....MRVRGI LRNC.PQWWI WG.ILGFWML MTCN.MEGNL WVTVYYGVPV
98BWMO37_d  ....MRVRGI LRNY.QQWWI WG.ILGFWML MICN.VVGNL WVTVYYGVPV
99BW3932_1  ....MRVRGI PRNW.QQWWI WS.ILG.... .FCS.VVGQL WVTVYYGVPV
99BW4642_4  ....MRVKGI LRNW.QQWWI WG.ILGFWML MICN.VVGNL WVTVYYGVPV
99BW4745_8  ....MRVRGI LRDY.QQWWI WS.ILGFWM. .ICN.GMGNL WVTVYYGVPV
99BW4754_7  ....MRVMGI KRNC.QQWWI WG.ILGFWML MI...CNGNL WVTVYYGVPV
99BWMC16_8  ....MRVMEI WRNC.PPWWI WG.ILGFWML MICN.GG.NR WVTVYYGVPV
A2_CD_97CD  ....TRVMGT QRNC.QKWWE WG.ILVFGMI MMCK.AAD.L WVTVYYGVPV
A2_CY_94CY  ....MRVMGT QRNY.QHLWR GG.ILILGML IMCK.ATD.L WVTVYYGVPV
A2D___97KR  ....MRVRGI QRNY.QHLWK WG.ILILGML MISK.ATEDL WVTVYYGVPV
A2G_CD_97C  ....MRVKGM QRNW.QNLWK WG.ALILGLV IICS.ASNNL WVTVYYGVPV
A_BY_97BL0  ....MKARGM XRNY.QHLWR XG.TMLFWMI IMCK.AAEDL VX.VYYXVPV
A_KE_Q23_A  ....MRVMGI QRNC.QHLLT WG.IMILGTI IFCS.AVENL WVTVYYGVPV
A_SE_SE659  ....MRVMGI QRNC.QHLLR WG.TIILGLI IICS.VADKL WVTVYYGVPV
A_SE_SE725  ....MRVMGT QMNW.QHLLR WG.TIILGMI MICS.TADNL WVTVYYGVPV
A_SE_SE753  ....MRAMGI QRNC.QHLLR WG.TMILGLV IICS.VAGNL WVTVYYGVPV
A_SE_SE853  ....MRVKGI QRNS.QHLLR WG.TMILGMI IICS.TADKL WVTVYYGVPV
A_SE_SE889  ....MRVMGT QMNW.QNLWR WG.TMILGII IICS.AAENL WVTVYYGIPV
A_SE_UGSE8  ....MRVMGT QRNC.QHLLN WG.IMILGMI IICS.TAENL WVTVYYGVPV
A_UG_92UG0  ....MRVMGI ERNY.PCWWT WG.IMILGMI IICN.TAENL WVTVYYGVPI
A_UG_U455_  ....MRVMGI QRNY.PCLWR WG.TMILGLI IICN..AQQL WVTVYYGVPV
AC_IN_2130  ....MRVRGI LRNY.QQWWI WG.SLGFWML MVCN.VVGNL WVTVYYGVPV
AC_RW_92RW  ....MRVMGT LMNY.QNLWG WG.TMILGML TICS.AANNL WVTVYYGVPV
AC_SE_SE94  ....MRVTGT QRSC.QPWWI WG.ILGFWML IICS.ATDKL WVTVYYGVPV
ACD_SE_SE8  ....MRVMGI QRNW.QHLLR WG.TMILGMI LICS.AVDKL WVTVYYGVPV
ACG_BE_VI1  ....MRVKGI QRNY.QQWWT WG.SLGLWML LICN.VMGNL WVTVYYGVPV
AD_SE_SE69  ....MRVRGI EMN.YQNLWR WG.TLLLGML MT.CSVTGRL WVTVYYGVPV
```

```
AD_SE_SE71  ....MRVMGI QRNC.QNLLT WG.TMILGMI IICS.VAENL WVTVYYGVPV
ADHK_NO_97  ....MKVMGT QRN.YPNWWR WG.VLILGML LICS.TTGNL WVTVYYGVPV
ADK_CD_MAL  ....MRVREI QRN.YQNWWR WG.MMLLGML MT.CSIAEDL WVTVYYGVPV
AG_BE_VI11  ....MRVRGT QMSWP.HLWN GG.ILILGLV IICS.ASNNL WVTVYYGVPV
AG_NG_92NG  ....MRVKGT QRNWQ.HLWT WW.TLILGLV IICS.ASNNL WVTVYYGVPV
AGHU_GA_VI  ....MRVMET QRN.YPRLWR WG.TIILGML MICN.AKENL WITVYYGVPV
AGU_CD_Z32  ....MKVKGI QRNC.QHLWK WG.TFILGLV IICS.AAENL WVTVYYGVPV
AJ_BW_BW21  ....MRVMET LMNCT.NLWR WG.LMIFGML MTCS.ATGNM WVTVYYGVPV
B_AU_VH_AF  ....MKVKET KRN.WQRLWR WG.IMLLGML MICS.ATEKL WVTVYYGVPV
B_CN_RL42_  ....MRVTGI RKN.YQHLWR WG.TMLLGML MICN.AAENL WVTVYYGVPV
B_DE_D31_U  ....MKVKEI RKN.YQHLWR WG.TMLLGML MICS.ATEKL WVTVYYGVPV
B_DE_HAN_U  ....MKVKET RKN.YQRLWR GG.TLLLGML MISS.VAGNL WVTVYYGVPV
B_FR_HXB2_  ....MRVKEK YQHLWRWGWR WG.TMLLGML MICS.ATEKL WVTVYYGVPV
B_GA_OYI__  ....MTARGT RKN.YQRLWR WG.TMLLGML MICS.AAENL WVTVYYGVPV
B_GB_CAM1_  ....MRAKGI RKN.CQRLWR WG.TMLLGML MICS.AADKL WVTVYYGVPV
B_GB_GB8_C  ....MKAKGT RKN.YQHLWK WG.IMLLGML MICS.ATEKL WVTVYYGVPV
B_GB_MANC_  ....MKVKEI RKN.YQNLWR WG.TLFLGML MICS.AEEKL WVTVYYGVPV
B_KR_WK_AF  ....MRVKGI RKN.YQHWWR WG.IMLLGMW MICS.AAEKL WVTVYYGVPV
B_NL_3202A  ....MKVKET RKN.YQHLWR WG.TMLLGML MICS.AAEQL WVTVYYGVPV
B_TW_TWCYS  ....MRVRGT RMN.CQHLWR WG.TMLLGML MISS.AAENL WVTVYYGVPV
B_US_BC_L0  ....MRVKEI RKN.YQHLWR WG.TMLFGIL MIYS.AAGNL WVTVYYGVPV
B_US_DH123  ....MRVMGI RKN.YQHLWK GG.TLLLGIL MICS.AAEQL WVTVYYGVPV
B_US_JRCSF  ....MRVKGI RKN.YQHLWK GG.ILLLGTL MICS.AVEKL WVTVYYGVPV
B_US_MNCG_  ....MRVKGI RRN.YQHWWG WG.TMLLGLL MICS.ATEKL WVTVYYGVPV
B_US_P896_  ....MRVKEI RKN.WQHLR. GG.ILLLGML MICSAAKEKT WVTIYYGVPV
B_US_RF_M1  ....MRVMEM RKN.CQHLWK WG.TMLLGML MICS.AAEDL WVTVYYGVPV
B_US_SF2_K  ....MKVKGT RRN.YQHLWR WG.TLLLGML MICS.ATEKL WVTVYYGVPV
B_US_WEAU1  ....MRVKGI RKN.YQHLWK WG.IMLLGIL MICS.AAENL WVTVYYGVPV
B_US_WR27_  ....MRVKGI RKN.CQHLWR WG.IMLLGML MICN.ATEQL WVTVYYGVPV
B_US_YU2_M  ....MRATEI RKN.YQHLWK GG.TLLLGML MICS.AAEQL WVTVYYGVPV
BF1_BR_93B  ....MRVRGM QRN.WQHLGK WG.LLFLGIL IICN..AENL WVTVYYGVPV
C_BR_92BR0  ....MRVEGI QRNW.KQWWI WG.ILGFWMV MIYN.VRGNL WVTVYYGVPV
C_BW_96BW0  ....MRVMGI QRNC.QQWWI WG.ILGFWMI INGM.GSW.. VT.VYYGVPV
C_BW_96BW1  ....MRVMGI MRNC.QPWWI WG.ILGFWML MICN.VMGKS WVTVYYGVPV
C_BW_96BW1  ....MRVRGI LRNY.LQWWI WG.ILGFWML MVCS.K.ENM WVTVYG.VPV
C_BW_96BW1  ....MRVRGI LRSW.QQWWI WG.TLGFW.. .ICS.GLGNL WVTVYDGVPV
C_ET_ETH22  ....MKVMGI QRNC.QQWWI WG.ILGFWML MICN.GMGNL WVTVYYGVPV
C_IN_93IN1  ....MRVRGT LRNY.QQWWI WG.VLGFWML MICN.GGGNL WVTVYYGVPV
C_IN_93IN9  ....MRVRGT LRNY.QQWWI WG.VLGFWML MICN.VGGNL WVTVYYGVPV
C_IN_93IN9  ....MRVRGI LRNY.QQWWI WG.ILGFWML MICN.VVGNL WVTVYYGVPV
C_IN_94IN1  ....MRVRGT LRNY.QQWWI WG.VLGFWML MICN.GGKDL WVTVYYGVPV
C_IN_95IN2  ....MRVRGI LRNY.QQWWI WG.VLGFWML MICN.VVGNL WVTVYYGVPV
CRF01_AE_C  ....MGVKGT QMNW.PHLWK WG.TLILGLV IICS.ASDTL WVTVYYGVPV
CRF01_AE_C  ....MRVKGT RRNW.PNLWK WG.TLILGLV IICS.ASDNL WVTVYYGVPV
CRF01_AE_C  ....MRVKGT QMNW.PNLWK WG.TLILGLV IMCS.ASDNL WVTVYG.VPV
CRF01_AE_T  ....MRVKET QMNW.PNLWK WG.TLILGLV IICS.ASDNL WVTVYYGVPV
CRF01_AE_T  ....MRVKET QMN..PNLWK WG.TLILGLV IICS.ASDDL WVTVYYGVPV
CRF01_AE_T  ....MRVKET QINW.PNLWK WG.TLILGLV IMCS.ASNNL WVTVYYGVPV
CRF01_AE_T  ....MRVKET QMSW.PNLWK WR.TLILGLV IICS.ASDNL WVTVYYGVPV
CRF01_AE_T  ....MRVKET QMNW.PNLWK WG.TLILGLV IICS.ASDNL WVTVYYGVPV
CRF01_AE_T  ....MRVKET QMNW.PNLWK WG.TLILGLV IICS.ASENL WVTVYYGVPV
CRF02_AG_F  ....MRVMGM QRNY.PLLWK WG.TIIFWIM IICN..AEKL WVTVYYGVPV
CRF02_AG_F  ....MRVMGI QRNY.PLFWK WG.MIIFWIM IICN..AEKL WVTVYYGVPV
CRF02_AG_G  ....MRVRGM QRNC.QNLWR WA..HDFWIL IICN.AAENL WVTVYYGVPV
CRF02_AG_N  ....MRVMGI QKNY.PLLWR WG.TNIFWIM IICN..AEQL WVTVYYGVPV
CRF02_AG_S  ....MRVMGI QKNY.PLLWR WG.MIIFWIM TICS..AGNL WVTVYYGVPV
CRF02_AG_S  ....MRVMGI LKSC.PPFWR WGMIMLLWIL IICN..AENL WVTVYYGVPV
CRF03_AB_R  ....MRVKEI RKH....LWR WG.TLFLGML MICS.ATENL WVTVYYGVPV
CRF03_AB_R  ....MRVKEI RKH....LWR WG.TLLLGML MICS.ATENL WVTVYYGVPV
CRF04_cpx_  ....MRVMGM QRN.YPHLWE WG.TLILGLV IICS.ASNNL WVTVYYGVPV
CRF04_cpx_  ....MRVMGI QRN.YPHLWE WG.TLILGLV IMCS.ASKDM WVTVYYGVPV
```

```
CRF04_cpx_   ....MTVMGT QRN.CPRLWT WG.TFILWLV IICS.ASNNL WVTVYYGVPV
CRF05_DF_B   ....MRVRGM QRN.WPHLGK WG.LLFLGIL IICS.ATDKF WVTVYYGVPV
CRF05_DF_B   ....MRVRGM QRN.WQHLGK WG.LLFLGIL IICS.AADNL WVTVYYGVPV
CRF06_cpx_   ....MRVKGI QTSWQ.HLWK WG.TLILGLV IICS.ASKNM WVTVYYGVPA
CRF06_cpx_   ....MRVRGI QKNWQ.HLWK WG.TLILGLV IICS.ASNNL WVTVYYGVPA
CRF06_cpx_   ....MRVKGI QMNWQ.HLWK WG.TLILGLA IICS.ATSNL WVTVYYGVPV
CRF06_cpx_   ....MTVKGI QRNWQ.HLWK WG.TLILGLV IICS.ASQNM WVTVYYGVPA
CRF11_cpx_   ....MRARGT QKNWH.DLWR WG.LMISGML MICN.ATDNL WVTVYYGVPV
CRF11_cpx_   ....MRVKET QRNWH.NLWR WG.LMIFGML MICN.AEK.M WVTVYYGVPV
D_CD_84ZR0   ....MRVKGI KRN.YQPLWK WG.IMLLGML MMTYSAADNL WVTVYYGVPV
D_CD_ELI_K   ....MRARGI ERN.CQNWWK WG.IMLLGIL MT.CSAADNL WVTVYYGVPV
D_CD_NDK_M   ....MRAREK ERN.CQNLWK WG.IMLLGML MT.CSAAEDL WVTVYYGVPI
D_UG_94UG1   ....MRVRET KRN.YQHLWK WG.TMLLGML MI.CSVTGKS WVTVYYGVPV
F1_BE_VI85   ....MRVRGM QRN.WQHLGK WG.LLFLGIL IICN.AADNL WVTVYYGVPV
F1_BR_93BR   ....MRVRGM QRN.WQHLGK WG.LLFLGTL IICN.AAENL WVTVYYGVPV
F1_FI_FIN9   ....MRVRGM QRN.WQHLGK WG.LLFLGML IICK.AADDL WVTIYYGVPV
F1_FR_MP41   ....MRVRVM QRN.WQHLGK WG.LLFLGIL IICS.AADNL WVTVYYGVPV
F2_CM_MP25   ....MRVREM QRN.WQHLGR WG.LLFLGIL IICS.AADKL WVTVYYGVPV
F2KU_BE_VI   ....MRVRER RRN.WQPLGR WG.ILFLGIF IICN.AAEDL WVTVYYGVPV
G_BE_DRCBL   ....MRVKGI QRNWQ.HLWN WG.ILILGLV IICS.AEK.L WVTVYYGVPV
G_NG_92NG0   ....MRVKGI QRNWQ.HLWK WG.TLILGLV IICS.ASDNL WVTVYYGVPV
G_SE_SE616   ....MRVTGI QRNW..HLWK WG.TLILGLV IICS.ASNNL WVTVYYGVPV
H_BE_VI991   ....TRVMET QRN.YPSLWR WG.TLILGML LICS.VVGNL WVTVYYGVPV
H_BE_VI997   .......TRV MRN.YPQWWR GG.ILLLGML LIYS.AAGNL WVTVYYGVPV
H_CF_90CF0   ....TRVMET QRN.YPSLWR WG.TLILGML LICS.AAQNL WVTVYYGVPV
J_SE_SE702   ....TRVMET QTSWL.SLWR WG.LMIFGML MICS.ARENL WVTVYYGVPV
J_SE_SE788   ....TRVMET QKNWQ.TLWR GG.LMIFGML MICK.AKEDL WVTVYYGVPV
K_CD_EQTB1   ....MRAREI QRN.WQHLGK RG.ILFLGIL IICS.AANNL WVTVYYGVPV
K_CM_MP535   ....MRVRGM QRN.WQTLGN WG.ILFLGIL IICS.NADKL WVTVYYGVPV
N_CM_YBF30   ....MKVMGM QSGWMGMKSG WLLFYLLVSL IKVIG.SEQH WVTVYYGVPV
O_CM_ANT70   ....MKAMEK RNK...K..L WTLYLAMALI TPCLSLR.QL YATVYAGVPV
O_CM_MVP51   ....MKVMKK NNR...K..S WSLYIAMALL IPCLSYSKQL YATVYSGVPV
O_SN_99SE_   MTVTMKVMEQ RNR...K..L GILCIVMALI TPCLSYN.QH YATVYAGVPV
O_SN_99SE_   MTVTMKVMEK RNR...K..L GILCMVMALI TPCLSHN.QH YATVYAGVPV
U_CD___83C   ....MRVKEI QRN.YQHLWK WS.LIILGMI MICK.AIEKS WVTVYYGVPV
```

```
                51                                                  100
00BW0762_1  WREAKTTLFC ASDAKAYDRE VHNVWATHAC VPTDPNPQEL VLENVTENFN
00BW0768_2  WKEAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEI VLGNVTENFN
00BW0874_2  WKEAKTTLFC ASDAKAYERE VHNVWATHAC VPTDPDPQEM VLENVTENFN
00BW1471_2  WREAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEM YLVNVTENFN
00BW1616_2  WKEAKTTLFC ASDAKAYDRE VHNVWATHAC VPTDPNPQEI GLENVTENFN
00BW1686_8  WKEAKTTLFC ASDAKAYEKE VHNIWATHAC VPTDPNPQEI VLENVTENFN
00BW1759_3  WRETKTTLFC ASDAKAYDKE VHNVWATHAC VPTDPNPQEL VLGNVTESFN
00BW1773_2  WKEAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEI PLKNVTENFN
00BW1783_5  WKEAKTTLFC ASDAKAYEKE AHNIWATHAC VPTDPNPREM FLENVTQNFN
00BW1795_6  WREAKAPLFC ASDAKAYDRE VHNVWATHAC VPTDPNPQEM VLKNVTENFN
00BW1811_3  WKEAKTTLFC ASDAKGYDRE VHNVWATHAC VPTDPNPQEL VLGNVTENFN
00BW1859_5  WKEAKTTLFC ASDAKAYERE VHNVWATHAC VPTDPNPQEI VLENVTENFN
00BW1880_2  WKEAKATLFC ASEAKAYESE VHNVWATHAC VPTDPNPQEI VLENVTENFN
00BW1921_1  WKEAKTTLFC ASDAKAYETE VHNVWATHAC VPTDPNPQEM ALENVTENFN
00BW2036_1  WREAKTTLFC ASDAKAYETE VHNIWATHAC VPTDPNPQEI VLGNVTENFN
00BW2063_6  WREAKATLFC ASDAKAYERE VHNVWATHAC VPTDPNPQEI VLENVTENFN
00BW2087_2  WKEAKTTLFC ASDAKAYERE VHNVWATHAC VPTDPNPQEM ELKNVTENFN
00BW2127_2  WKEAKAPLFC ASDAKAYEKE AHNVWATHAC VPTDPNPQEI ELKNVTENFN
00BW2128_3  WKEAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEM VLENVTENFN
00BW2276_7  WKEAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEL VLENVTENFN
00BW3819_3  WREAKATLFC ASDAKAHERE VHNVWATHAC VPTDPNPQEM VMENVTENFN
00BW3842_8  WKEAKTTLFC ASDAKGYETE VHNVWATHAC VPTDPDPQEM VLGNVTENFN
00BW3871_3  WREAKTTLFC ASDAKAYERE VHNVWATHAC VPTDPNPQEM LLKNVTENFN
00BW3876_9  WKEAKTTLFC ASDAKVYEKE VHNVWATHAC VPTDPNPQEM VLDNVTENFN
00BW3886_8  WKEAKTTLFC ASDAKAYERE VHNVWATHAC VPTDPNPQEI VLGGVTENFN
00BW3891_6  WREAKTTLFC ASDAKGYEKE VHNVWATHAC VPTDPDPQEM VLENVTENFN
00BW3970_2  WKEAKTTLFC ASDAKGYERE VHNIWATHAC VPTDPNPQEM FLHNVTENFN
00BW5031_1  WKEAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPSPQEI VLENVTETFN
 96BW01B21  WKEAKTTLFC ASDAKAYEKG VHNVWATHAC VPTDPNPQEV FLENVTENFN
  96BW0407  WKEAKATLFC ASDARAYDRE VHNVWATHAC VPTDPNPQEV ILENVTENFN
  96BW0502  WKEAKTTLFC TSDAKAYETE VHNVWATHAC VPTDPNPQEI VLENVTENFN
 96BW06_J4  WKEAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPSPQEI VLGNVTENFN
 96BW11_06  WREAKATLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEV VLENVTENFN
  96BW1210  WKEAKTTLFC ASDAKAYEGE VHNVWATHAC VPTDPNPQEL VLGNVTENFN
 96BW15B03  WREASNTLFC ASYAKAYEKE VHNVWATHAC VPTDPNPQEI ELDNVTENFN
 96BW16_26  WKEAKTTLFC ASDAKAYDKE VHNVWATHAC VPTDPNPQEI ILKNVTENFN
 96BW17A09  WREAKTTLFC ASDAKAFESE VHNVWATHAC VPTDPNPQEM VLENVTENFN
 96BWMO1_5  WREAKTTLFC ASDAKAYEAE VHNVWATHAC VPTDPNPQEI ELKNVTENFN
 96BWMO3_2  WKEAKTTLFC ASDAKAYERE VHNVWATHAC VPTDPNPQEM VLENVTENFN
98BWMC12_2  WREAKTTLFC ASDAKAYERE VHNIWATHAC VPTDPNPQEM VLENVTENFN
98BWMC13_4  WKEAKAPLFC ASDAKVYEKE VHNVWATHAC VPTDPNPQEL VLENVTENFN
98BWMC14_a  WREATTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEI VLGNVTENFN
98BWMO14_1  WKEAKTTLFC ASDAKAYDKE VHNVWATHAC VPTDPSPQEM FLANVTENFN
98BWMO18_d  WREAKATLFC ASNAKAYEKE VHNVWATHAC VPTDPNPQEM VLENVTENFN
98BWMO36_a  WKEAKATLFC ASDAKAYDKE VHNVWATHAC VPTDPDPQEI VLENVTESFN
98BWMO37_d  WKEAKTTLFC ASDAKAYDKE VHNVWATHAC VPTDPNPQEM VLENVTENFN
99BW3932_1  WKEAKATLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQKL VLGNVTENFN
99BW4642_4  WKEAKTTLFC ASDAKAYDKE VHNVWATHAC VPTDPNPQEI VLENVTENFN
99BW4745_8  WREAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEL VLKNVTENFN
99BW4754_7  WREAKTTLFC ASDAKAYDKE VHNVWATHAC VPTDPNPQEI VLENVTENFN
99BWMC16_8  WREAKATLFC ASDAKAYERE VHNVWATHAC VPTDPDPQEI ALENVTENFN
A2_CD_97CD  WRDADTTLFC ASDAKAYATE KHNVWATHAC VPTDPNPQEV NLANVTEDFN
A2_CY_94CY  WKDADTILFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEI NLENVTENFN
A2D___97KR  WRDAETTLFC ASDAKAYDTE AHNVWATHAC VPTDPNPQEI NLENVTENFN
A2G_CD_97C  WEDANTPLFC ASDAKSYSTE RHNVWATHAC VPTDPNPQEM ILENVTESFN
A_BY_97BL0  XXDAATTLFC ASDAKAXDKE VHNVWATHAC VPTDPDPQEI ILGNVTEKFD
A_KE_Q23_A  WRDADTTLFC ASDAKAYETE KHNVWATHAC VPTDPNPQEI HLDNVTEKFN
A_SE_SE659  WKDAETTLFC ASDAKAYDPE VHNVWATHAC VPTDPNPQEM HLENVTEESN
A_SE_SE725  WKDAETTLFC ASDAQAYKTE MHNVWATHAC VPTDPNPQEL HLKNVTEEFN
A_SE_SE753  WKDAETTLFC ASDAKAYETE KHNVWATHAC VPTDPNPQEI HLKNVTEKFN
```

```
A_SE_SE853   WKDAETTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEL HLANVTEEFN
A_SE_SE889   WRDAETTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEI NLENVTEEFN
A_SE_UGSE8   WKDAETTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEL YLENVTEDFN
A_UG_92UG0   WKDANTTLFC ASDAKAYDTE VHNVWATHAC VPTDPSPQEL KMENVTEEFN
A_UG_U455_   WKDAVTTLFC ASDAKAYDAE VHNVWATHAC VPTDPNPQEI DLVNVTEEFN
AC_IN_2130   WKDAETTLFC ASDAKAYETE KHNVWATHAC VPTDPNPQEI HLENVTEDFN
AC_RW_92RW   WKDAETTLFC ASDAKAYDPE KHNVWATHAC VPIDPDPQEI HLENVTEEFN
AC_SE_SE94   WKEAKTTLFC ASDAKAYEAE VHNVWATHAC VPTDPNPHEI NLENVTENFN
ACD_SE_SE8   WKDAETTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV LLGNVTEDFN
ACG_BE_VI1   WKEAKTTLFC ASDAKAYETE AHNVWATHAC VPTDPSPQEI ELKNVTENFN
AD_SE_SE69   WRDAETTLFC ASDAKAYDAE VHNVWATHAC VPTDPNPQEI NLENVTEEFN
AD_SE_SE71   WKDAETTLFC ASDAKAYETE KHNVWATHAC VPTDPNPQEI HLENVTENFN
ADHK_NO_97   WKEAKTTLFC ASDAKAYDTE MHNVWATHAC VPTDPSPQEI VLENVTENFN
ADK_CD_MAL   WKEATTTLFC ASDAKSYETE VHNIWATHAC VPTDPNPQEI ELENVTEGFN
AG_BE_VI11   WEDADTTLFC ASDAKAYSTE SHNVWATHAC VPTDSNPQEI PLENVTENFN
AG_NG_92NG   WEDADTPLFC ASDAKAYSTE RHNVWATHAC VPTDPNPQEI TLENVTETFN
AGHU_GA_VI   WRDAKTSLFC ASDAKSYSTE SHNVWATHAC VPTDPNPQEI NLENVTENFN
AGU_CD_Z32   WKDAETTLFC ASDAKAYDTE KHNVWATHAC VPTDPNPQEL SLGNVTEKFN
AJ_BW_BW21   WKEAKTTLFC ASNAKAYSTE GHNIWATHAC VPTDPNPQEI ILENVTENFN
B_AU_VH_AF   WKEATTTLFC ASDAKAYDKE VHNVWATHAC VPTDPNPQEI LLENVTEEFN
B_CN_RL42_   WKEATTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV VLGNVTENFN
B_DE_D31_U   WKEATTTLFC ASDAKAYDKE VHNVWATHAC VPTDPDPQEV VLENVTEDFN
B_DE_HAN_U   WKEATTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV VMGNVTENFN
B_FR_HXB2_   WKEATTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV VLVNVTENFN
B_GA_OYI__   WKEATTTLFC ASDARAYATE VHNVWATHAC VPTDPNPQEV VLGNVTENFD
B_GB_CAM1_   WKEATTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV VLENVTENFN
B_GB_GB8_C   WKEATTTLFC ASDAKAYDTE KHNVWATHAC VPTDPNPQEV VLGNVTENFN
B_GB_MANC_   WKEATTTLFC ASDAKAHGTE VHNVWATHAC VPTDPNPQEI VLENVTEYFD
B_KR_WK_AF   WKEATTTLFC ASDAKAYDTE VHNVWATHAC VPIDPNPQEV FLENVTENFN
B_NL_3202A   WKEATTTLFC ASDAKASDTE VHNVWATHAC VPTDPNPQEV ALENVTEDFD
B_TW_TWCYS   WKEATTTLFC ASDAKTYDTE VHNVWATHAC VPTDPNPQEV TLGNVTENFN
B_US_BC_L0   WKEATTTLFC ASDAKAYDTE VHNVWATHAC IPTDPNPQEI VLENVTEDFN
B_US_DH123   WKEANTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEI LLENVTEDFN
B_US_JRCSF   WKETTTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV VLENVTEDFN
B_US_MNCG_   WKEATTTLFC ASDAKAYDTE VHNVWATQAC VPTDPNPQEV ELVNVTENFN
B_US_P896_   WREATTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV VLGNVTENFN
B_US_RF_M1   WKEATTTLFC ASEAKAYKTE VHNVWAKHAC VPTDPNPQEV LLENVTENFN
B_US_SF2_K   WKEATTTLFC ASDARAYDTE VHNVWATHAC VPTDPNPQEV VLGNVTENFN
B_US_WEAU1   WKEATTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV VLENVTENFN
B_US_WR27_   WKEATTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV RLENVTEDFN
B_US_YU2_M   WKEATTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEV KLENVTENFN
BF1_BR_93B   WKEATTTLFC ASDAKAYEKE AHNVWATHAC VPTDPNPQEV VLENVTENFD
C_BR_92BR0   WKEAKTTLFC ASDAKAYDAE VHNVWATHAC VPTDPNPQEM VLENVTENFN
C_BW_96BW0   WKEAKATLFC ASDARAYDRE VHNVWATHAC VPTDPNPQEV NLENVTENFN
C_BW_96BW1   WREAKATLFC ASDAKAYERE VHNVWATHAC VPTDPNPQEV VLENVTENFN
C_BW_96BW1   WKEAKTTLFC ASDAKAYEGE VHNVWATHAC VPTDPNPQEL VLGNVTENFN
C_BW_96BW1   WREASNTLFC ASYAKAYEKE VHNVWATHAC VPTDPNPQEI ELDNVTENFN
C_ET_ETH22   WKDASPTLFC ASDAKAYDTE VHNVWGTFAC VPTDPSPQEL GLENVTENFN
C_IN_93IN1   WKEAKTTLLC ASDAKAYERE VHNVWATHAC VPTDPNPQEI VLGNVTENFN
C_IN_93IN9   WKEAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEI PLGNVTENFN
C_IN_93IN9   WKEAKTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEM VLENVTENFN
C_IN_94IN1   WKEAKTTLFC ASDAKAYGKE VHNVWATHAC VPTDPNPQEI SLENVTENFN
C_IN_95IN2   WKEANTTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEI VMENVTENFN
CRF01_AE_C   WRDADTTLFC ASDANAQETE VHNVWATHVC VPTDPNPQEI HLENVTENFD
CRF01_AE_C   WRDADTILFC ASDAKAHVTE VHNVWATHAC VPTDPNPQEI YLENVTENFD
CRF01_AE_C   WRDADTTLFC ASDAKAHETE VHNIWATHAC VPTDPNPQEI DLENVTENFN
CRF01_AE_T   WKDADTTLFC ASDAKAHETE VHNVWTTHAC VPTDP.PQEI HLENVTENFN
CRF01_AE_T   WRDADTTLFC ASDAKAHETE VHNVWATHAC VPTDPNPQEI HLENVTENFN
CRF01_AE_T   WRDADTTLFC ASDAKAHETE VHNVWATHAC IPTDPNPQEM HLENVTENFN
CRF01_AE_T   WRDADTTLFC ASDAKAHEAE VHNVWATHAC VPTDPNPQEI HLENVTENFN
CRF01_AE_T   WRDADTTLFC ASDAKAHETE VHNVWATHAC VPTDPNPQEI HLENVTENFN
```

```
CRF01_AE_T  WRDADTTLFC ASDAKAHETE VHNVWATHAC VPTDPNPQEI HMENVTENFN
CRF02_AG_F  WRNAETTLFC ASDAKAYDAE VHNVWATHAC VPTDPNPQEI HLKNVTEKFN
CRF02_AG_F  WRDAETTLFC ASDAKAYDVE VHNVWATHAC VPTDPNPQEI HLKNVTEKFN
CRF02_AG_G  WKTADTTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEI HLENVTEEFN
CRF02_AG_N  WKTAETTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEI HLENVTEKFN
CRF02_AG_S  WRDAETTLFC ASDAKAYDTE VHNVWATHAC VPTDPNPQEI HLGNVTEDFN
CRF02_AG_S  RRDAETTLFC ASDAKAYETE VHNVWATHAC VPTDPSPQEM HLENVTENFN
CRF03_AB_R  WKEATTTLFC ASDAKAYSKE VHNVWATYAC VPTDPSPQEI PLKNVTENFN
CRF03_AB_R  WKEATTTLFC ASDAKAYSKE VHNVWATYAC VPTDPSPQEI PLENVTENFN
CRF04_cpx_  WRDAETTLFC ASEAKAYEKE VHNIWATHAC VPTDPNPQEV ALINVTENFN
CRF04_cpx_  WRDAETKPXC ASDAKAYDKE IHNIWATHAC VPTDPNPQEL ALTNVTENFN
CRF04_cpx_  WRDAETSPFC ASDAKAYDKE VHNIWATHAC VPTDPNPQEI FLKNVTEDFN
CRF05_DF_B  WREAKTTLFC ASDAKGYDKE VHNVWATHAC VPTDPNPQEV VLGNVTENFN
CRF05_DF_B  WKEANTTLFC ASDAKGYEKE AHNVWATHAC VPTDPNPQEL ALENVTENFN
CRF06_cpx_  WEDADTILFC ASDAKAYSPD KHNVWATHAC VPTDPNPQEI SLKNVTENFN
CRF06_cpx_  WEDADTILFC ASDAKAYSAE KHNVWATHAC VPTDPNPQEI PLKNVTENFN
CRF06_cpx_  WEDADTILFC ASDAKAYSAE KHNVWATHAC VPTDPNPQEI PLENVTENFN
CRF06_cpx_  WEDADTILFC ASDAKAYSAE KHNVWATHAC VPTDPNPQEI KLENVTENFN
CRF11_cpx_  WKDANTTLFC ASDAQAYSPE KHNVWATHYC VPTDPNPQEI LLGNVTENFN
CRF11_cpx_  WRDADTILFC ASDARTYSTE KHNVWATHSC VPTDPNPREL SLENVTENFN
D_CD_84ZR0  WKEATTTLFC ASDAKSYKTE AHNIWATHAC VPTDPNPQEI ELKNVTENFN
D_CD_ELI_K  WKEATTTLFC ASDAKSYETE AHNIWATHAC VPTDPNPQEI ALENVTENFN
D_CD_NDK_M  WKEATTTLFC ASDAKAYKKE AHNIWATHAC VPTDPNPQEI ELENVTENFN
D_UG_94UG1  WKEATTTLFC ASDAKAYKAE AHNIWATHAC VPTDPNPQEI KLENVTENFN
F1_BE_VI85  WKEATTTLFC ASDAKAYERE AHNVWATHAC VPTDPNPQEV FLKNVTENFD
F1_BR_93BR  WKEATTTLFC ASDAKSYEKE AHNVWATHAC VPTDPNPQEV VLENVTERFN
F1_FI_FIN9  WKEANTTLFC ASDAKSYEKE VHNVWATHAC VPTDPNPQEV ALN.VTENFN
F1_FR_MP41  WKEATTTLFC ASDAKGYERE VHNVWATHAC VPTDPNPQEI WLKNVTENFD
F2_CM_MP25  WKEATTTLFC ASDAKAYERE VHNVWATYAC VPTDPSPQEL VLGNVSEKFN
F2KU_BE_VI  WKEANTTLFC ASDAKPYDTE VHNVWVTHAC VPTDPNPQEV FLQNVTENFN
G_BE_DRCBL  WEDANAPLFC ASDAKAHSTE SHNIWATHAC VPTDPSPQEI NMRNVTENFN
G_NG_92NG0  WEDADTPLFC ASDAKSYSSE KHNVWATHAC VPTDPNPQEI AIENVTENFN
G_SE_SE616  WEDADTTLFC ASDAKSYSAE SHNVWATHAC VPTDPNPQEI IMENVTEYFN
H_BE_VI991  WKEAKTTLFC ASDAKAYDTE RHNVWATHAC VPTDPNPQEM VLENVTETFN
H_BE_VI997  WKEAKTTLFC ASDAKAYEPE KHNVWATHAC VPTDPSPQEM VLANVTENFN
H_CF_90CF0  WKEAKTTLFC ASDAKAYETE KHNVWATHAC VPTDPNPQEM VMENVTESFN
J_SE_SE702  WRDAKTTLFC ASDAKAYSTE KHNVWATHAC VPTDPNPQEM SLPNVTENFN
J_SE_SE788  WKDAKTTLFC ASDAKAYSTE KHNVWATHAC VPTDPSPQEM NLPNVTENFN
K_CD_EQTB1  WKEATTTLFC ASDAKAYETE VHNVWATHAC VPTDPNPQEV VLENVTENFN
K_CM_MP535  WKEATPTLFC ASDAKAYEKE VHNVWATHAC VPTDPNPQEV EMENVTENFN
N_CM_YBF30  WREAETTLFC ASDAKAHSTE AHNIWATQAC VPTDPNPQEV LLPNVTEKFN
O_CM_ANT70  WEDATPVLFC ASDANLTSTE KHNIWASQAC VPTDPTPYEY PLHNVTDDFN
O_CM_MVP51  WEEAAPVLFC ASDANLTSTE QHNIWASQAC VPTDPNPHEF PLGNVTDNFD
O_SN_99SE_  WEEATPVLFC ASDANLTSTE QHNIWASQAC VPTDPSPYEY PLTKVTDNFN
O_SN_99SE_  WEEATPVLFC ASDVNLTSTE QHNIWASQAC VPTDPSPYEY PLKNVTDNFN
U_CD___83C  WKDAETTLFC ASDAKAYEKE SHNVWATHAC VPTDPSPQEL VLGNVTENFN
```

```
                101                                                  150
00BW0762_1  MWKNYMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLNCINV TN........
00BW0768_2  MWKNDMVDQM HEDVISLWDQ SLKPCVKLTP LCITLNCTSI NG........
00BW0874_2  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTNY NN........
00BW1471_2  MWKNDMVDQM HEDIISIWDQ SLKPCVKLTP LCVTLYCTNV TKR.......
00BW1616_2  MWKNDMVNQM HEDIISLWDQ SLKPCVRLTP LCVTLNCNNV TT........
00BW1686_8  MWKNDVVDQM HEDIISLWDQ SLKPCVKLTP LCVTLQCSNV S.........
00BW1759_3  MWKNDMVDQM HEDIISLWDQ SLNPCVKLTP LCVTLKCSNV N.........
00BW1773_2  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLECEDA N.........
00BW1783_5  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLTCSNL NI........
00BW1795_6  MWKNDMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLDCTNA TI........
00BW1811_3  MWKNDMVDQM HEDIISLWDE SLKPCVKLTP LCVTLNCINA TN........
00BW1859_5  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLTCKDY N.........
00BW1880_2  MWENGMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLNCSNA KA........
00BW1921_1  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLKCSDA K.........
00BW2036_1  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTNA NV........
00BW2063_6  MWENDMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLKCRNA NN........
00BW2087_2  MWKNDMVDQM HEDIISLWDE SLKPCVKLTP LCVTLNCVTV NCT.......
00BW2127_2  MWKNDMVEQM HEDIIRLWDE SLKPCVRLTP LCVTLRCSNA GSG.......
00BW2128_3  MWKNDMVDQM HEDIISLWDE SLKPCVKLTP LCVTLSCNAT N.........
00BW2276_7  MWKNDMVDQM HEDIISLWDE SIKPCVKLTP LCVTLNCSIV N.........
00BW3819_3  MWKNDMVDQM HEDIISLWDE SLKPCVKLTQ LCVTLECSNV N.........
00BW3842_8  MWENDMVDQM HEDVISLWDE SLKPCVKLTP LCVTLNCTNY NG........
00BW3871_3  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLMCSNV T.........
00BW3876_9  MWENDMVDQM HEDIISLWDQ SPKPCVKLTP LCVTLKCTDA T.........
00BW3886_8  MWKNDMVDQM HEDVISLWDE SLKPCVKLTP LCVTLKCGNV NN........
00BW3891_6  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLKCVSI N.........
00BW3970_2  MWKNDMVDQM HEDIISLWDE SLKPCVKLTP LCVTLECKNV TTN.......
00BW5031_1  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLNCSAA NN........
 96BW01B21  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLKCSTY N.........
  96BW0407  MWKNDMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTKV NGT.......
  96BW0502  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLKCRNV N.........
 96BW06_J4  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTS LCVTLHCSNV N.........
 96BW11_06  MWENDMVNQM HEDIISLWDQ SLKPCVKLTP LCVTLNCT.. ..........
  96BW1210  MWKNDMVDQM HEDIISLWDE SLKPCVKLTP LCVTLNCSNN VTR.......
 96BW15B03  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLKCTNY ST........
 96BW16_26  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLDCTNV T.........
 96BW17A09  MWKNDMVDQM HEDIISLRDQ SLKPCVKLTP LCVTLNCTNA TN........
 96BWMO1_5  MWENDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLNCKNV TS........
 96BWMO3_2  MWKNDMADQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTEA KV........
98BWMC12_2  MWKNDMVDQM HEDIIRLWDQ SLKPCVKMTP LCVTLNCANF NAS.......
98BWMC13_4  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLKCSNV TV........
98BWMC14_a  MWRNDMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLDCTNV T.........
98BWMO14_1  MWENDMVDQM HQDIISLWDE SLKPCVKLTP LCVTLNCRNA NLN.......
98BWMO18_d  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLTCTNA TKNVTN....
98BWMO36_a  MWKNDMVDQM HEDVISIWDQ SLKPCVKLTP LCVTLNCSNV N.........
98BWMO37_d  MRDNDMVDQM HEDIINLWDQ SLKPCVRLTP LCVTLNCKDA SVN.......
99BW3932_1  MWKNDMVDQM HEDMIRLWDQ SLKPCVKLTP LCVTLKCREV N.........
99BW4642_4  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLKCTNV N.........
99BW4745_8  MWKNDMVDQM HEDVISLWDQ SLKPCVKLTP LCVTLICSNN I.........
99BW4754_7  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLNCNKV TV........
99BWMC16_8  MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLNCVNV TKNVTK....
A2_CD_97CD  MWKNNMVEQM HADIISLWDQ SLKPCVKLTP LCVTLNCSNA NTTNT.....
A2_CY_94CY  MWKNNMVEQM QEDIISLWDQ SLKPCVKLTP LCVILNCSNA NTSTH.....
A2D___97KR  MWKNGMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCSRV KNTIS.....
A2G_CD_97C  MWKNDMVEQM HVDIISLWDQ SLKPCVKLTP FCVTLNCTNA TFPNA.....
A_BY_97BL0  MXKNNXVEQM QTDIISL.DQ SLKPCVKLTP LCVTLNCAEP NSTRS.....
A_KE_Q23_A  MWKNNMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLHCTNV TSV.......
A_SE_SE659  MWKNNMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLNCTNV ..........
A_SE_SE725  MWKNSMVEQM HTDIISLWDE SLKPCVKLTP LCVTLNCTNA ..........
A_SE_SE753  MWKNYMVEQM HTDIISLWDQ SLEPCVKLTP LCVTLECHYN ITV.......
```

```
A_SE_SE853   MWKNSMVEQM HTDIISLWDQ SLIPCVKLTP LCVTLECNDY NYN.......
A_SE_SE889   MWKNNMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLNCSSV TN........
A_SE_UGSE8   MWKNNMVEQM HTDIISLWDQ SLKPCVQLTP LCVTLNCSNN VTA.......
A_UG_92UG0   MWKNNMVEQM HTDIISLWDQ SLKPCVQLTP LCVTLDCSYN ITN.......
A_UG_U455_   MWKNNMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLDCHNI TIN.......
AC_IN_2130   MWKNSMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLNCSNV NG........
AC_RW_92RW   MWKNNMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLECNNI TNVNN.....
AC_SE_SE94   IWKNYMVEQM HQDIISLWDQ SLKPCVKLTP LCVTLNCRDA TV........
ACD_SE_SE8   MWKNNMVEQM HTDIISLWDQ SLQPCVKLTP LCVTLNCTNV TIT.......
ACG_BE_VI1   MWKNDMVDQM HQDIISLWDE SLKPCVKLTP LCVTLNCSNV TAIN......
AD_SE_SE69   MWKNNMVEQM HTDIISLWDQ SLKPCVQLTP LCVTLNCNNV TNKIN.....
AD_SE_SE71   MWKNNMVKQM HTDIISLWDQ SLQPCVKLTP LCVTLHCNDT ..N.......
ADHK_NO_97   MWENNMVDQM HTDIISLWDQ SLKPCVKLTP LCVTLNCTDP AN........
ADK_CD_MAL   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTNV NGTAVNG.TN
AG_BE_VI11   MWKNNMVDQM HEDIISLWDE SLKPCVKLTP LCVTLTCTNV NCTNN.....
AG_NG_92NG   MWKNNMVEQM HEDIISLWDE SLKPCVKLTP LCVTLNCTNV NCNSN...VT
AGHU_GA_VI   MWKNNMVEQM HTDIISLWDQ SLKPCVQITP LCVTLECSKI N.........
AGU_CD_Z32   MWKNNMVEQM HEDVISLWDQ SLKPCVKLTP LCVTLSCSDI R.........
AJ_BW_BW21   IWKNDMVEQM QEDIISVWDE SLKPCVKLTP LCVTLNCTNA TVSNT.....
B_AU_VH_AF   MWKNNMVEQM HEDIISLWDQ SLKPCVQLTP LCVTLNCTDE LT........
B_CN_RL42_   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTNL K.........
B_DE_D31_U   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTDL K.........
B_DE_HAN_U   MWKNDMVEQM QEDIISLWDQ SLKPCVKLTP LCVTLKCTDY N.........
B_FR_HXB2_   MWKNDMVEQM HEDIISLWDQ SLKPCVKLTP LCVSLKCTDL K.........
B_GA_OYI__   MWKNNMVEQM QEDIISLWDQ SLKPCVKLTP LCVTLDCTDV NTTSSS....
B_GB_CAM1_   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLICTNV NN........
B_GB_GB8_C   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTDL R.........
B_GB_MANC_   MWKNNMVEQM QEDVISLWDQ SLKPCVKLTP LCVTLDCTDY VG........
B_KR_WK_AF   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLKCTDL NDTNTN....
B_NL_3202A   MWKNNMVEQM HEDIINLWDQ SLKPCVKLTP LCVTLNCTDF G.........
B_TW_TWCYS   MWKNNMADQM QEDIISLWDE SLKPCVELTP LCVTLKCNDT ..........
B_US_BC_L0   MWKNNMVEQM QEDIISLWDQ SLKPCVKLTP LCVTLNCTDE LKNA......
B_US_DH123   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLHCTDL K.........
B_US_JRCSF   MWKNNMVEQM QEDVINLWDQ SLKPCVKLTP LCVTLNCKDV ..........
B_US_MNCG_   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTDL R.........
B_US_P896_   MWKNNMVDQM HEDIISLWDE SLKPCVKLTP LCVTLNCTNL ..........
B_US_RF_M1   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTDA NLN.......
B_US_SF2_K   MWKNNMVEQM QEDIISLWDQ SLKPCVKLTP LCVTLNCTDL G.........
B_US_WEAU1   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTNV NVTN......
B_US_WR27_   MWKNNMXEQM HEDIIXLWDQ SLKPCVKLTP LCVTLNCTDV ..........
B_US_YU2_M   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTDL ..........
BF1_BR_93B   MWKNNMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLRCSNA TT........
C_BR_92BR0   MWENDMVEQM HQDIISLWDQ SLKPCVKLTP LCVTLHCSNR T.........
C_BW_96BW0   MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTEV NGTSDSS...
C_BW_96BW1   MWENDMVNQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTNV TV........
C_BW_96BW1   MWKNDMVDQM HEDIISLWDE SLKPCVKLTP LCVTLNCSNN VTR.......
C_BW_96BW1   MWKNDMVDQM HEDIISLWDQ SLKPCVKLTP LCVTLKCTNY ST........
C_ET_ETH22   MWKNDMVEQM HQDIISLWDQ GLKPCVKLTP LCVTLNCNAI KNNTKVT...
C_IN_93IN1   MWKNDMVDQM HEDVISLWDQ SLKPCVKLTP LCVTLECRNV S.........
C_IN_93IN9   MWKNDMVNQM HEDVISLWDQ SLKPCVKLTP LCVTLECKNV K.........
C_IN_93IN9   MWKNDMVNQM HEDVISLWDQ SLKPCVKLTP LCVTLECSEY NGTSKAN...
C_IN_94IN1   MWKSDMVNQM HEDVISLWDQ SLKPCVKLTP LCVTLECGNV T.........
C_IN_95IN2   MWKNDMVNQM HEDVISLWDQ SLKPCVKLTP LCVTLECRNV NST.......
CRF01_AE_C   MWKNNMVEQM QEDVISLWDQ SLKPCVKLTP LCVTLHCTKA KLNDTYN...
CRF01_AE_C   MWKNNMVEQM QEDVISLWDQ SLQPCVKLTP LCVTLHCTKA SFTNATS...
CRF01_AE_C   MWKNNMVEQM QEDVISL.DQ SLKPCVKLTP LCVTLDCTKA DFYTTKF...
CRF01_AE_T   MWKNNMVEQM QEDVISLWDQ SLQPCVKLTP LCVTLHCTTA KLTNVTN...
CRF01_AE_T   MWKNNMVEQM QEDVISLWDQ SLKPCVKLTP LCVTLNCTNA NLTNVNN...
CRF01_AE_T   MWKNNMVEQM QEDVISLWDQ SLKPCVKLTP LCVTLNCTNA NWTNANV...
CRF01_AE_T   MWKNNMVEQM QEDVISLWDQ SLKPCVKLTP LCVTLNCTTA NFTNFNL...
```

```
CRF01_AE_T   MWKNKMAEQM QEDVISLWDQ SLKPCVKLTP LCVTLNCTNV NATNVSN...
CRF02_AG_F   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLDCHNV NSS.......
CRF02_AG_F   MWKNNMVEQM HEDIISLWDQ SLKPCVELTP LCVTLDCYNV SS........
CRF02_AG_G   MWKNNMVEQM HVDIISLWDQ SLKPCVKLTP LCVTLDCQNF KN........
CRF02_AG_N   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLDCHNF NN........
CRF02_AG_S   MWKNSMVEQM HEDIISLWDQ SLKPCVQLTP LCVTLHCQDN LT........
CRF02_AG_S   MWKNNMVEQM HVDIISLWDQ SLKPCVKLTP LCVTLECHNY NYT.......
CRF03_AB_R   MGKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTDL KK........
CRF03_AB_R   MGKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTEV KT........
CRF04_cpx_   MWKNDMVEQM HEDIISLWNE GLKPCAKLTS LCVTFTCINA T.........
CRF04_cpx_   MWENSTVEQM HEDIISLWDE GLKPCVKLTP LCVALNCSNA TIIINS....
CRF04_cpx_   MWKNNMVEQM HEDIISLWEE GLKPCVKLTP LCVALNCGDA TIK.......
CRF05_DF_B   MWKNDMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLNCTDF KA........
CRF05_DF_B   MWKNNMVEQM HADIISLWDQ SLKSCVKLTP LCVTLNCTDA TS........
CRF06_cpx_   MWKNNMVDQM HEDIISLWDE SLKPCVKLTP LCVTLTCTNA TLGNKTLGNN
CRF06_cpx_   MWENHMVEQM HEDIISLWDE SLKPCVKLTP LCVTLICTNI NITSTNS...
CRF06_cpx_   MWKNHMVQQM HEDIISLWDE SLKPCVKLTP LCVTLNCTNV TDHGIN....
CRF06_cpx_   MWKNNMVEQM HEDIISLWEE SLKPCVKLTP LCVTLNCTNV NAT......K
CRF11_cpx_   MWKNNMVEQM HEDIISLWDE SLKPCVKLTP LCVTLNCAEV TS........
CRF11_cpx_   MWKNNMVEQM HEDVISLWDE SLKPCVKLTP LCVALNCTDA R.........
D_CD_84ZR0   MWKNNMVDQM HEDIISLWDQ SLKPCVKLTP RCVTLNCTDA SRN......S
D_CD_ELI_K   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCSDE LRNNG....T
D_CD_NDK_M   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTDE LRN......S
D_UG_94UG1   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLNCTNW VTD.......
F1_BE_VI85   MWKNNMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLNCTNA TN........
F1_BR_93BR   MWENNMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLDCRNI AT........
F1_FI_FIN9   MWENDMVEQM HKDIISLWDQ SLKPCVKLTP LCVTLNCTNA TT........
F1_FR_MP41   MWKNNMVEQM HEDIISLWDQ SLKPCVKLTP LCVTLHCSDV NI........
F2_CM_MP25   MWKNNMVDQM HEDIISLWDE SLKPCVKLTP LCVTLNCTKA II........
F2KU_BE_VI   MWKNNMVEQM HADIISLWDQ GLQPCVKLTP LCVTLNCSEK IN........
G_BE_DRCBL   MWKNNMVEQM HEDIISLWDE SLKPCVKLTP LCVTLNCTEI N...N.....
G_NG_92NG0   MWKNNMVEQM QEDIISLWEE SLKPCVKLTP LCITLNCTNV N.........
G_SE_SE616   MWKNNMVEQM HEDIISLWDE SLKPCVKLTP LCVTLNCTDV TNKGNKR.NN
H_BE_VI991   MWVNDMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLDCSSV NA........
H_BE_VI997   MWDNDMVEQM QTDIISLWDQ SLKPCVKLTP LCVTLDCSNI TR........
H_CF_90CF0   MWENNMVEQM HTDIISLWDQ SLKPCVKLTP LCVTLNCTNV RN........
J_SE_SE702   MWKNDMVDQM QEDIISVWDE SLKPCVKITP LCVTLNCSDV NSNNS.....
J_SE_SE788   MWKNDMVDQM QEDIISVWDE SLKPCVKITP LCVTLNCSNI TSNSN.....
K_CD_EQTB1   MWKNNMVEQM HTDIISLWDE SLKPCVKLTP LCVTLTCTNV TN........
K_CM_MP535   MWKNNMVEQM HTDIISLWDE SLKPCVELTP LCVTLNCTDY KG........
N_CM_YBF30   MWENKMADQM QEDIISLWEQ SLKPCVKLTP LCVTMLCNDS YGEER.....
O_CM_ANT70   IWKNYMVEQM QEDIISLWDQ SLKPCVQMTF LCVQMECTN. ..........
O_CM_MVP51   IWKNYMVDQM HEDIISLWEQ SLKPCEKMTF LCVQMNCVD. ..........
O_SN_99SE_   IWKNYMVEQM QEDIISLWEQ SLKPCVQMTF LCVQMNCTNY VQ........
O_SN_99SE_   IWENYMVEQM QEDIISLWEQ SLKPCVQMTF LCVQMNCTN. ..........
U_CD___83C   MWKNKMVEQM HEDIISLWDQ SLKPCVKLTF LCVTLNCIDV KN........
```

```
                151                                                
```

```
A_SE_SE853   ...VTNSSHS Y..... NVTNMQ.... EMKNCSFNVT TELRDKRQKV
A_SE_SE889   ......SSVT N..... ITSDMAG... EIKNCSFNMT TEIRDKRQKV
A_SE_UGSE8   ...NTNSTSA N..... LTDSVKG... EMRNCSFNIT TELRDKKKKV
A_UG_92UG0   ...NITNSIT N..... SSVNMRE... EIKNCSFNMT TELRDKNRKV
A_UG_U455_   ...NTN.NNT N..... ITDGVR...E EMKNCSFNMT TELRDKKQKV
AC_IN_2130   ...NSTGWGK ..... .....E EIKNCSFNIT TELRDKRQKV
AC_RW_92RW   ...TVN.... ..... ITDDMKG... EIKNCSFNMT TELRDKKQRV
AC_SE_SE94   ...TPNNATH N..... DSM..V...G DMKNCPFNMT TELRDKRRKE
ACD_SE_SE8   ...TNATDSN N..... ..ASLQDMAK EMTNCSFNMT TELRDKKQRV
ACG_BE_VI1   ....SNGTAI N..... ITESIKG... EMKNCSFKAT TEIKDKKKKE
AD_SE_SE69   .....ETSMN G..... ..... EIKNCSFNMT TELRDKEQQV
AD_SE_SE71   ...VTNATNI T..... NANTITG... EMKNCSFNMT TELMDKKRKV
ADHK_NO_97   ...HTDTTNN ..... ...TSIQPSQ PSANCSFNVT TAIRDKQQKV
ADK_CD_MAL   AGSNRTNAEL KM..... .....EIG EVKNCSFNIT PVGSDKR.QE
AG_BE_VI11   .STREIRGKN CSLD..... .....TEVG ELKNCSFNIT TELRDKKKTE
AG_NG_92NG   STGNSAGTNA TCNI..... .....EEAN NLKNCSFNIT TEIRDKKKTE.
AGHU_GA_VI   ...ITNNSTD KANV..... ...TNN..DA EMRNCSFNIT TEIRDRKKKE
AGU_CD_Z32   .....NSTES N..... ITAEMQG... EIKNCSYNMT TELRDKQRKI
AJ_BW_BW21   .....GCTNN NCT..... .....VS EMKECHFNIT GGGR..RKKE
B_AU_VH_AF   ...NVTFTNS RHVTNS.... .SYVGSMEKG EMKNCSFNIT TSIRDKRHKE
B_CN_RL42_   ...NATNTSS T..... .....MEGG EIKNCSFNIT TSIKTKVK.D
B_DE_D31_U   ...NATNTNN SSWT..... ...MTGEMKG EIKNCSFNIT TSIRDKVQKE
B_DE_HAN_U   ...NATNSS. ..... ...WGRMEKG EIQNCSFKVT TNIRDKVQKE
B_FR_HXB2_   ...NDTNTNS SS.G..... ...RMIMEKG EIKNCSFNIS TSIRGKVQKE
B_GA_OYI__   .LRNATNTTS SS..... ...WETMEKG ELKNCSFNTT TSIRDKMQEQ
B_GB_CAM1_   ...TRTNSSD WDRR..... ...EGEKMKG EIKNCSFNVT TSIRNKVRKE
B_GB_GB8_C   ...NDTNTNN SIME..... .....GG EMKNCSFNIT TSIRDKMQKE
B_GB_MANC_   ...NATNTTS TNNTAS.... .GSWGAMR.G EIKNCSFNIT TNIRDKVHKE
B_KR_WK_AF   .N.SSTSENN TNPTIS.... ..GGEGMGEG EMKNCSFNVT TNIRDKVQKE
B_NL_3202A   ...NATNTTS S..S..... ...GVIIEKG EIKNCSFKIN TNMKDKAQIE
B_TW_TWCYS   ...TMSKNDS N..... ...TLTMEKG EIKNCSFNVT TSLRNKVQKE
B_US_BC_L0   .TNTTSTNTP SGS..... ...WKKMERG EIKNCSFNVL G...DKKQKA
B_US_DH123   ...NGTNLKN GTK..... ..IIGKSMRG EIKNCSFNVT KNIIDKVKKE
B_US_JRCSF   ...NATNTTS ..... .SSEGMMERG EIKNCSFNIT KSIRDKVQKE
B_US_MNCG_   ...NTTNTNN STANN.NS.. .NSEGTIKGG EMKNCSFNIT TSIRDKMQKE
B_US_P896_   ...NITKNTT N....PTS.. .SSWGMMEKG EIKNCSFYIT TSIRNKVKKE
B_US_RF_M1   ....GTNVTS SSG..... ...GTMMENG EIKNCSFQVT TSRRDKTQKK
B_US_SF2_K   ...KATNTNS SN..... ...WKEEIKG EIKNCSFNIT TSIRDKIQKE
B_US_WEAU1   .LKNETNTNS SSG..... ...GEKMEEG EMKNCSFNVT TLIRNKRKTE
B_US_WR27_   ...WNATSTS KNTTITNS.. .SNERPMEKG EMKNCSFSIT TSIRDKVQKE
B_US_YU2_M   .R.NATNTTS SS..... ...WETMEKG EIKNCSFNIT TSIRDKVQKE
BF1_BR_93B   ...NST.... .QND..... ...TLKEEPG AIQNCSFNMT TEVRDKQLKV
C_BR_92BR0   ....IDYN.. N..... RTDNMGG... EIKNCSFNMT TEVRDKREKV
C_BW_96BW0   V..PANGTSN ..... SSVSMKE... EMRNCSFNIT TELRDKNKQE
C_BW_96BW1   .NDTLHQNFT D..... ..... .MKNCSFNVT TELRDKRKHE
C_BW_96BW1   ....NS..NA T..... YNNKNNG... EIKNCSFNAT TEIRDKQQKV
C_BW_96BW1   .NYSNTMNAT S..... YNNNTTE... EIKNCTFNMT TELRDKKQQV
C_ET_ETH22   .....N N..... SINSAND... EMKNCSFNIT TELRDKKRKA
C_IN_93IN1   ....RNVSSY N..... TYNGSVE... EIKNCSFNAT PEVRDRKQRM
C_IN_93IN9   ....NDSTHN E..... TYTESVK... EIKNCSFNAT TEIRDRKQTV
C_IN_93IN9   ATNNVNATSN G..... NATSNGE... EIQQCFFNVT TEMRDKKQRV
C_IN_94IN1   ....QNGTYN D..... ...ESNK... EITNCTFNTT TEIRGRKQKV
C_IN_95IN2   ....GNGTHS K..... TYNESMK... EIKNCSFNAT TVIKDKKQTV
CRF01_AE_C   GTAKL....N ..... .....DTIGD EVRNCSFNVT TELRDKKQEV
CRF01_AE_C   DRIK..... ..... .....MED AVRNCSFNMT TELQDKKQEV
CRF01_AE_C   NTTEK....P E..... ..IEISEMQK EVSNCPFNIT TELRDKEQEV
CRF01_AE_T   I.....TNVP N..... ..IG..NITD EVRNCSFNMT TEIRDKKQKV
CRF01_AE_T   .......NVS N..... ..IIG.NITD EVRNCSFNMT TELRDKKQQV
CRF01_AE_T   TNVN..NNVT N..... ..IVG.NITE EVRNCSFNMT TELIDKKQKV
CRF01_AE_T   TKADNMTNVS N..... ..ITIGNITD EVRNCTFNMT TDLIDKKQKV
CRF01_AE_T   K.....TNVS N..... ..IIG.NITD EVRNCTFNMT TELTDKKQKV
```

```
CRF01_AE_T  T.....TEAP N..........  ..IV...GTD EVKNCSFNVT TELRDKTQQV
CRF02_AG_F  ...NSSTSNS SNSSTPINRT  IDSDMQE...  EIKNCSFNMT TELRDKKQKV
CRF02_AG_F  ...NSSISVK S.........  ISRDMQG...  EIKNCSFNMT TELRDKKQQV
CRF02_AG_G  ..........  ..........  ISDGSNS...  EIKNCSFNMT TELRDKKQKV
CRF02_AG_N  ....SYSNSS N.........  LTSDMNG...  EIKNCSFNIT TEVRDKKKKM
CRF02_AG_S  ....SS.GN.  ..........  ISENMQG...  EIKNCSFNMT TELRDKKQKV
CRF02_AG_S  ...RNN....  ..........  SKINEVQ...  EMKNCSFNMT TVLKDKKKKM
CRF03_AB_R  ...EVTSTNT S.........  .....SIKMM  EMKNCSFNIT TDLRDKVKKE
CRF03_AB_R  ...NDTSTNA S.........  .....GIEMM  KN..CSFNIT TDLRDKVKKE
CRF04_cpx_  ....TTNSTN GTVI......  ...K.....E  GIKNCSFDIT TEIRDKKKKE
CRF04_cpx_  TNNSTTNSTG NSTV......  ...KS...TA  EIKNCSFNIT TEVRDKQKKE
CRF04_cpx_  ....STNSTP NVTT......  ...T.....N  EMKNCSFNIT TEIRDKKKKA
CRF05_DF_B  ...NSTANST TNST......  ...TLKEETG  AVQNCSFNMT TEVNDKKLKV
CRF05_DF_B  ...ATTTSKN ISATPTSN..  PNDTLKEEQG  AIQNCTFNIT TEVKDKNKRV
CRF06_cpx_  STNSTLGNNS TIVD......  ......DISK  EIKNCSFNIT TEIRDKTKKE
CRF06_cpx_  .....T.KNI TVES......  ........GE  EIKNCSFNVT TEIRDKQKEE
CRF06_cpx_  .......NNN TVEG......  ........KE  EIKNCSFNVT TEIKDKKKKE
CRF06_cpx_  YSNETVGKSL TVKD......  ........RE  EIKNCSFNIT TEVRDQKKTE
CRF11_cpx_  ..........  ..........  ......YNTT  EMKNCSFNVT TELIDRRKQE
CRF11_cpx_  ..........  ..........  ......DNAT  DIKNCTFNIT TELEDKKKNE
D_CD_84ZR0  TDNNSTLPTV KP........  .......GE.  .MKNCSFNIT TVVTDKRKQV
D_CD_ELI_K  MGNNVTTEEK G.........  ..........  .MKNCSFNVT TVLKDKKQQV
D_CD_NDK_M  KGNGKVEEEE K.........  ..........  .RKNCSFNVR ....DKREQV
D_UG_94UG1  .....TTNTT G.........  ..........  .MANCSFNIT TEIRDKKKQV
F1_BE_VI85  ...NSQ....  ..........  ...EK...PG  AMQNCSFNMT TEVRDKKLKL
F1_BR_93BR  ...NGTNDTI AIND......  ...TLKEDPE  AIQNCSFNTT TEIRDKQLKV
F1_FI_FIN9  ...TNDTLS. DQSS......  ...TLKEEPG  AIQNCSFNMT TEVEDKKQKV
F1_FR_MP41  ...TSNATTT NDTS......  ...TP.EESG  AIQNCSFNMT TEVKDKKLRV
F2_CM_MP25  ...NVTSSNN TTLA......  ...PNVTISE  EMKNCSFNIT TEIRDKQKKE
F2KU_BE_VI  ...INSTDLT NWANKTNNWA  NETTLLNITT  GMRNCSFNIT TMLKDKKKKQ
G_BE_DRCBL  ........NS TRNI......  ......TEEY  RMTNCSFNMT TELRDKKKAE
G_NG_92NG0  .SANHTEANN ..TV......  ......ENKE  EIKNCSFKIT TERGGKKKEE
G_SE_SE616  STDNSTETNN S.TV......  ......DNPG  EIKNCSFNVT TEIRDKKKKE
H_BE_VI991  ...TNVTKSN NSTD......  ...INIGEIQ  EQRNCSFNVT TAIRDKNQKV
H_BE_VI997  ...NDTNSSS TVNA......  ...TSSPSAN  ELTNCSFNVT TVIRDKQQRV
H_CF_90CF0  ...NTSNSTS SMEA......  ...GG.....  ELTNCSFNVT TVLRDKQQKV
J_SE_SE702  .....TDSNS SASN......  ......NSPE  IMKNCSFNVT TEIRNKRKQE
J_SE_SE788  .....TTSNS SVS.......  .......SPD  IMTNCSFNIT TEIRNKRKQE
K_CD_EQTB1  ...NRTNANK NDT.......  NINATVTSTD  EIKNCSFNIT TELKDKKKRV
K_CM_MP535  ...TNSTN..  ..........  NATSTVVSPA  EIKNCSFNIT TEIKDKKKKE
N_CM_YBF30  ...NNTNMTT R.........  ...EPDIGYK  QMKNCSFNAT TELTDKKKQV
O_CM_ANT70  ....IAG...  ..........  .....TTNEN  LMKKCEFNVT TVIKDKKEKK
O_CM_MVP51  ...LQTNKTG  ..........  LLN...ETIN  EMRNCSFNVT TVLTDKKEQK
O_SN_99SE_  ..GNYTNNSS  ..........  INNDTSSPEN  LVKQCEFNVT TVVKDKKEKK
O_SN_99SE_  ..VNDETNSS  ..........  VKNDTSSSEN  LMKKCEFNVT TVLKDKKEKQ
U_CD___83C  ......STNN N.........  ......TEEA  TITNCSFKVP TELKDKTETV
```

	201				250
00BW0762_1	YALFYRLDIV	QLGE.....	.....NNAN	SE.....	.....YRLI
00BW0768_2	HALFYRLDIV	PLDEKDK...	.....SSN	SN.....	.....YRLI
00BW0874_2	SALFYRLDIV	PLNGS.....	.....ERNK	SE.....	.....YRLI
00BW1471_2	RALFYRLDIV	PLNESDN...	.....NSY	RE.....	.....YRLI
00BW1616_2	YAIFHSLDIV	PLEN.....	.....SEN	SE.....	.....YRLI
00BW1686_8	YALFYKLDIV	PLEE.....	.....NDI	ST.....	.....YRLI
00BW1759_3	HALFYRLDIV	PLEGE.....	.....NNTN	NE.....	.....YRLI
00BW1773_2	HALFYRLDIV	QLD.....	.....N	SS.....	.....YRLI
00BW1783_5	YALFYKLDIV	PLEGNNS...	.....E..	.....	.....YRLI
00BW1795_6	YALFYRLDIV	SLDNENN...	.....KT.	AE.....	.....YRLI
00BW1811_3	YALFYKPDIV	PLDGS.....	.....NS	SE.....	.....YRLI
00BW1859_5	YALFYKIDIV	PLND....N.	.....NSN.N	SM.....	.....YRLI
00BW1880_2	YALFYRLDVV	PLDSPS....	.....NATN	SR.....	.....YRLI
00BW1921_1	YALFYRLDVV	QLN.....	.....	SE.....	.....YRLI
00BW2036_1	YALFYKLDIV	PLNGNSG...	.....	SE.....	.....YRLI
00BW2063_6	YALFYKLDIV	PLGNTNG...	.....T..	.E.....	.....YRLI
00BW2087_2	YALFYKLDIV	SLDD.....	.....NN	S.....	.....YRLI
00BW2127_2	YALFYRLDVV	PLDND.....	.....SA	TN.....	.....YRLI
00BW2128_3	YALFYKLDIV	PLNNS.....	.....SDNSS	GE.....	.....YRLI
00BW2276_7	QALFYKLDIV	PLNSTGE...	.....NNN	TE.....	.....YRLI
00BW3819_3	YALFYRLDVV	PLNGK.....	.....NS	SS.....	.....YRLI
00BW3842_8	HALFYRLDIV	PLEDNSG...	.....NSS	SN.....	.....YRLI
00BW3871_3	YALFYKLDIV	PLND.....	.....N..	NE.....	.....YRLI
00BW3876_9	NALFYKLDVV	PLHE.....	.....GN	S.....	.....YRLI
00BW3886_8	YALFYRLDIV	PLHDSSS...	.....DG.	SE.....	.....YVLI
00BW3891_6	HALFYRLDIV	PLNG.....	.....KNQS	NE.....	.....YRLI
00BW3970_2	NALFYTLDIV	PLDENQ....	.....	N.....	.....YRLI
00BW5031_1	FALFNILDIV	PLNNEN....	.....NTKN	SD.....	.....YRLI
96BW01B21	YALFYKFDVV	PLN.....	.....GNNT	SE.....	.....YILI
96BW0407	RALFYSLDIV	QPNN.....	.....S	TE.....	.....YRLI
96BW0502	HALFYRLDVV	PLQG.....	.....NN	NE.....	.....YRLI
96BW06_J4	YALFYRLDVV	PLGD.....	.....N..	SS.....	.....YRLI
96BW11_06	YALFYRLDIV	PLNNKNE...	.....S..	SE.....	.....YRLI
96BW1210	YALFYRLDIV	PLDN.....	.....NS	SE.....	.....YRLI
96BW15B03	YALFYKLDIV	PLNSNS....	.....	SE.....	.....YRLI
96BW16_26	YALFYRLDVV	PLNGE.....	.....NSNSS	GE.....	.....YRLI
96BW17A09	SALFYRLDIV	PLNENNS...	.....SSN	SE.....	.....YRLI
96BWMO1_5	YALFYKLDIV	PLTNDAS...	.....EN.	SE.....	.....YRLI
96BWMO3_2	YALFYKLDIV	PLDGNNE...	.....DGN	KQ.....	.....YRWI
98BWMC12_2	SALFYRLDIV	PLK.....	.....ENS	SE.....	.....YRLI
98BWMC13_4	QALFYRLDIV	PLDNANG...	.....T..	SE.....	.....YRLI
98BWMC14_a	YALFYRLDIV	PLGE.....	.....D..	SS.....	.....YRLI
98BWMO14_1	YALFYKLDIV	ELDG.....	.....NS	SN.....	.....YVLI
98BWMO18_d	SALFYKLDIV	PLD.....	.....NSS	SK.....	.....YILI
98BWMO36_a	YVLFYKLDIV	PLNGNG....	.....SN	SE.....	.....YRLI
98BWMO37_d	YALFYRPDIV	PLNEG.....	.....	N.....	.....YRLI
99BW3932_1	YALFYRLDIV	PLKN.....	.....S	SE.....	.....YRLI
99BW4642_4	NALFYKLDIV	PLNE....K.	.....ANNSY	SY.....	.....YRLI
99BW4745_8	YALFYRIDIV	PLDE.....	.....NNNS	SE.....	.....YRLI
99BW4754_7	HALFYRLDIV	PLETK.....	.....NSNE	SA.....	.....YRLI
99BWMC16_8	YALFYKVVIV	PLSE.....	.....NST	SE.....	.....YRLI
A2_CD_97CD	YSLFYELDVV	LLNRSKN...	.....SSY	ST.....	.....YRLI
A2_CY_94CY	YSLFYRLDVV	QLDESENKNT	.....SGSN	TL.....	.....YRLI
A2D___97KR	QALFYELDIV	QLNSSDSND.	.....TLN	RQ.....	.....YRLI
A2G_CD_97C	RSLFYTLDIV	QINKDNN...	.....	.T.....	.....YRLI
A_BY_97BL0	HSLFYKLDIV	STSNNDSX..	.....	.Q.....	.....YRLI
A_KE_Q23_A	YSLFYRLDIV	PINEN.....	.QG.....	SE.....	.....YRLI
A_SE_SE659	HSLFYRLDIV	QMNEN.....	.RGNSSNSSY	NE.....	.....YRLI
A_SE_SE725	YSLFYKLDIV	QINDN.....	.GNNSNNS..	SE.....	.....YRLI
A_SE_SE753	YSLFYRLDLV	KIDEN.....	.KSN...SSN	SK.....	.....YRLI

```
A_SE_SE853   TSLFYKLDVV PIGGN..... .......DTNS TQ..... .......YRLI
A_SE_SE889   HALFYRLDVV PMDN..... .......NNS. .L..... .......YRLI
A_SE_UGSE8   YSLFYKLDIV KINKNKSFRG .KNSSGNSSS DR..... .......YRLI
A_UG_92UG0   YSLFYKLDVV QINNG..... .......NNSS NL..... .......YRLI
A_UG_U455_   YSLFYRLDIV QINKTD.... ..........N NS..... .......YRLI
AC_IN_2130   YSLFYRLDVV PIEEGQGNS. .....SNSGY KE..... .......YRLI
AC_RW_92RW   YSLFYRLDIV QINSNSN... .......NSSH NQ..... .......YRLI
AC_SE_SE94   HALFYRLDIV PLDEGNSNSN ....ESNNNY SD..... .......YRLI
ACD_SE_SE8   YSLFYKLDVV QINSN..... .......QNNS SQ..... .......YRLI
ACG_BE_VI1   YALFYRLDIL PLNKENK... ........GSS GK..... .......YRLI
AD_SE_SE69   HSLFYKLDVV QMGNSNS... .......... SQ..... .......YRLI
AD_SE_SE71   YSLFYKLDVV QINENQ.... ..YNSSNNSN KE..... .......YRLI
ADHK_NO_97   HALFYRVDLV SIDN..... .......NDNN TQ..... .......YRLI
ADK_CD_MAL   YATFYNLDLV QIDDSDN... ..........S S..... .......YRLI
AG_BE_VI11   HALFYRPDVV PINNDN.... .......... ..... ....SSYMLI
AG_NG_92NG   YALFYRLDVV PIDGNNNV.. .......... .S..... ....NNYRLI
AGHU_GA_VI   YALFYKHDLV PITN..... ........ET KT..... .......FILI
AGU_CD_Z32   YSLFYRLDIV PIEENSSNG. ......N..S SE..... .......YRLI
AJ_BW_BW21   YALFYKEDIA LIKDRPN... ........NS NY..... ....SEYILV
B_AU_VH_AF   FALFYKLDVV QIDGS..... ..........N TS..... .......YRLI
B_CN_RL42_   YALFYKVDVV PIGND..... ..........S TS..... .......YRLI
B_DE_D31_U   YAHFYKLDVV PIDND..... ..........N TS..... .......YRLI
B_DE_HAN_U   SALFYKTDVV PIDNNKTS.. .....NRDNT TS..... .......YMLI
B_FR_HXB2_   YAFFYKLDII PIDND..... ..........T TS..... .......YKLT
B_GA_OYI__   YALFYKLDVL PIDKN..... ..........D TK..... .......FRLI
B_GB_CAM1_   YALFYKLDVV PIDKAN.... .......... TS..... .......YTLI
B_GB_GB8_C   YALLYKLDIV SIGSD..... ..........N TS..... .......YILT
B_GB_MANC_   YALFYKLDVV PIEKK..... ..........N TS..... .......FRLI
B_KR_WK_AF   YALFYKLDII PIDN..... .......... TS..... .......YALR
B_NL_3202A   YALFYKLDVV PIDNN..N.. .....TNTSY TS..... .......YRLI
B_TW_TWCYS   YASFYRLDLV QTDEN..... ..........S TS..... .......YRLI
B_US_BC_L0   YALFYKLDVV PIDNDKNS.. .......... TK..... .......YRLI
B_US_DH123   YALFYRHDVV PIDRN..... ..........I TS..... .......YRLI
B_US_JRCSF   YALFYKLDVV PID..... .......NKNN TK..... .......YRLI
B_US_MNCG_   YALLYKLDIV SIDND..... ..........S TS..... .......YRLI
B_US_P896_   YALFNRLDVV PIE..... .......NTNN TK..... .......YRLI
B_US_RF_M1   YALFYKLDVV PIEKGNISPK N.NTSNNTSY GN..... .......YTLI
B_US_SF2_K   NALFRNLDVV PIDN..AS.. .....TTTNY TN..... .......YRLI
B_US_WEAU1   YALFYKLDVM PIDHDNTS.. .......... ..... .......YTLI
B_US_WR27_   HALFYRLDVV PIDK..... .......NNTN TS..... .......YRLI
B_US_YU2_M   YALFYNLDVV PIDN..... .......... AS..... .......YRLI
BF1_BR_93B   HALFYRLDIV PISNDNSSND .......NSS RE..... .......YRLI
C_BR_92BR0   HALFYRLDIV PLKNE..... .....SSNTS GD..... .......YRLI
C_BW_96BW0   RARFYRLDIV QLNNN..... .......SNS NE..... .......YRLI
C_BW_96BW1   YALFYRLDIV PLNNKNE... .......SN. SE..... .......YRLI
C_BW_96BW1   YALFYRLDIV PLDN..... ........NS SE..... .......YRLI
C_BW_96BW1   YALFYKLDIV PLNSNS.... .......... SE..... .......YRLI
C_ET_ETH22   YALFYKLDIV PLN..... .......NGS TD..... .......YRLI
C_IN_93IN1   YALFYGLDIV PLN...KKN. .....SSENS SE..... .......YRLI
C_IN_93IN9   YALFYRLDIV QLN...SDD. .....KKNSS EY..... .......YRLI
C_IN_93IN9   HALFYRLDLV PLDNENKSS. .....FSNSS KT..... .......YRLI
C_IN_94IN1   YALFYKLDIV PIS...ETS. .....NQS.. .......... .......RLI
C_IN_95IN2   YALFYKLDIV PLDNEEQEN. .....DSNSS GY..... .......YRLI
CRF01_AE_C   HALFYVPDIV RIG....EK. ..NKNSSGNS SE..... .......YILI
CRF01_AE_C   HALFYTSDVV QISSSVQNNN NSNTSGQNNS HK..... .......FRLI
CRF01_AE_C   YALFYRSDLV PIE..... ..RNSGENNG SS..... .......YRLI
CRF01_AE_T   HALFYKLDIV QIEDK..... .......NDS SK..... .......YGLI
CRF01_AE_T   HALFYKLDIV QMN..... .......KNS SE..... .......YRLI
CRF01_AE_T   YALFYKLDIR QMN..... .......SNS SE..... .......YRLI
CRF01_AE_T   YALFYKLDTI PIG..... .......NNN NM..... .......YRLI
CRF01_AE_T   HALFYKLDIV QIEDK..... .......KTS SE..... .......YRLI
```

```
CRF01_AE_T  QALFYKLDIV QMGG..... .......NDS GE..... ......YRLI
CRF02_AG_F  SALFYRLDVV QINES..... ........GN SQ..... ......YRLI
CRF02_AG_F  SALFYRLDVV QINES..... ........SN SQ..... ......YRLI
CRF02_AG_G  NALFYRVDVV QMNNS..... .......... .Q..... ......YRLI
CRF02_AG_N  HALFYRLDVV QINEN..... ........NG SQ..... ......YRLI
CRF02_AG_S  YALFYRYDVV QINETG.... ........DN IQ..... ......YRLI
CRF02_AG_S  AALFYKIDIV PIDKN..... .........A TY..... ......YRLI
CRF03_AB_R  YALFYKLDVV QIDND..... .......... .S..... ......YRLI
CRF03_AB_R  HALFYKLDVV QIDND..... .......... .S..... ......YRLI
CRF04_cpx_  YALFYRIDIV PINARVPING ...SNRNNST EE..... ......YMLI
CRF04_cpx_  HALFYRLDVV PINNNVPINN ...TSNTSEY RE..... ......YRLM
CRF04_cpx_  YALFYRLDIV PINDNNSTN. ...SRRSSNT SD..... ......YMLI
CRF05_DF_B  HALFYRLDIV PISSD....D .......SSN SS..... ......YRLI
CRF05_DF_B  HALFYRLDIV SINS..... .......SRK E..... ......YRLI
CRF06_cpx_  YALFYRPDIV PIGDD..... .........S NN..... ....SDYRLI
CRF06_cpx_  YALFYRLDVV PINDG..... .........S NN..... ....NSYRLI
CRF06_cpx_  RALFYTLDVV PINDN..... .........G NN..... ....STYRLI
CRF06_cpx_  YALFYRPDVM QVDG..... .........K NS..... ....STYRLI
CRF11_cpx_  YALFYKLDIV PINDNNN... .........SS NV..... ....SDYRLI
CRF11_cpx_  RALFYRLDVV PINDS..... .........SS NI..... ....GQYRLI
D_CD_84ZR0  HALFYRLDVV QIDNEGKNE. .....INDTY GT..... ......YRLI
D_CD_ELI_K  YALFYRLDIV PIDNDSS... .......TNS TN..... ......YRLI
D_CD_NDK_M  YALFYKLDIV PIDNNNR... .......TNS TN..... ......YRLI
D_UG_94UG1  QALFYKLDVV KINDNDS... .........DN TS..... ......YRLI
F1_BE_VI85  SALFYRLDIV PIGNN...N. .........S SE..... ......YRLI
F1_BR_93BR  HALFYKLDIV QINKD..DN. .......... RT..... ......YRLI
F1_FI_FIN9  HALFYRLDIE PISNN...N. .........SR EE..... ......YRLI
F1_FR_MP41  NALFYKLDII PINNS..... .........SS SD..... ......YRLI
F2_CM_MP25  YALFYKLDVV QINNS..... .......... NTS..... ......YRLI
F2KU_BE_VI  YALFYREDIV PINIKKNNKT .......NSN SKKNNNTSNN SIENSKYRLI
G_BE_DRCBL  YALFYRTDVV PINEMNNENN ........GT NS..... ....TWYRLT
G_NG_92NG0  YALFYKLDVV PISNGN.... .......... .K..... ....TSYRLI
G_SE_SE616  YAFFYRLDVV PINN..... .......... .A..... ....TNYRLT
H_BE_VI991  HALFYRADIV QIDEGER... ......NKSD NH..... ......YRLI
H_BE_VI997  HALFYRLDVV PIDETSNNN. ......NSNS TK..... ......YRLI
H_CF_90CF0  HALFYRLDVV PIDNNS.... .......... TQ..... ......YRLI
J_SE_SE702  YALFYRQDVV PIN..... .........S DN..... ....KSYILI
J_SE_SE788  YALFYRQDVV PID..... .........S NN..... ....KNYILI
K_CD_EQTB1  SALFYKLDIV QIKQSEINQS .......ESE .......... ......DRLI
K_CM_MP535  SALFYRLDVL PLN.GEGNNS .......STE .......... ......YRLI
N_CM_YBF30  YSLFYVEDVV PINAYN.... .......... .......... ....KTYRLI
O_CM_ANT70  QALFYVSDLM ELNETSSTNK .........T NS..... ....KMYTLT
O_CM_MVP51  QALFYVSDLS KVNDSNAVN. .......... .G..... ....TTYMLT
O_SN_99SE_  QALFYVSDLM KINEANDT.. .......... .K..... ....DMYTLI
O_SN_99SE_  QALFYVSDLM KVNENND... .......... .......... ....TMYTLI
U_CD___83C  HTLFYKLDVV PLNVTN.... .........N SS..... ..ISSTYRLI
```

```
           251                                                        300
00BW0762_1 NCNTSTITQA CPKVNFDPIP IHYCAPAGYA ILKCNTKTFD GTGPCTNVST
00BW0768_2 NCNTSAVTQA CPKVSFEPIP IHYCAPAGYA ILKCNNKTFN GTGPCHNVST
00BW0874_2 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
00BW1471_2 NCNTSTITQA CPKVTFDPIP IHYCAPAGYA ILKCNNETFN GTGPCNNVST
00BW1616_2 KCNTSTITQA CPKVNFDPIP IHYCAPAGYA ILKCRNKTFN GTGPCNNVST
00BW1686_8 NCNTSSISQA CPKVSFGPIP IHYCAPAGYA ILKCNNKTFN GTGPCQNVST
00BW1759_3 NCNTSAVTQA CPKVTFDPIP VHYCAPAGYA ILKCNNKTFN GAGPCNNVST
00BW1773_2 NCNTSAITQA CPKVSFDPIP IHYCTPAGYA ILKCNNQTFN GTGPCNDVSS
00BW1783_5 NCNTSAITQA CPKVSFEPIP IHFCAPAGYA ILKCNNKTFN GTGPCNNVST
00BW1795_6 NCNTSTITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GKGPCNNVST
00BW1811_3 NCNTSALTQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCTNVST
00BW1859_5 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCNDKTFN GTGPCQNVST
00BW1880_2 NCNTSAITQA CPKINFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
00BW1921_1 NCNTSAITQA CPKVTFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
00BW2036_1 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCQNVST
00BW2063_6 NCNTSTITQS CPKVTFDPIP IHYCAPAGYA ILKCNNETFN GTGPCNNVST
00BW2087_2 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVSI
00BW2127_2 NCNTSAITQA CPKISFDPIP IHYCAPTGYA ILKCNNKTFN GTGPCNNVST
00BW2128_3 NGNTSALTQA CPKVSFDPIP IHYCTPAGYA ILKCNNKTFN GTGPCNNVST
00BW2276_7 NCNTSAITQA CPKITFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCQNVSP
00BW3819_3 NCNTSAVTQS CPKISFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
00BW3842_8 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA IIKCNNKTFN GIGPCQNISI
00BW3871_3 NCNTSAISQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCQNVST
00BW3876_9 HCNTSTITQA CPKVSFEPIP IHYCAPAGYA ILKCNDKTFS GTGPCLNVST
00BW3886_8 NCNTSTITQA CPKVSFDPIP IHYCAPAGYA ILNCNNKTFN GTGPCQNVST
00BW3891_6 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
00BW3970_2 NCNTSKVTQA CPKVSFDPIP LHYCAPAGYA ILKCNNNTFN GTGPCNNVST
00BW5031_1 SCNTSTITQA CPKVSWDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
 96BW01B21 NCNTSAISQA CPKVSFDPIP IHYCTPAGFA ILKCNNRTFN GTGPCNNVST
  96BW0407 NCNTSTITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GLGPRNNVST
  96BW0502 NCNTSAITQA CPKVSFDPIP IHYCTPAGYA ILKCNNQTFN GTGPCNNVSS
 96BW06_J4 NCNTSTISQA CPKISFDPIL IHYCAPGGYA ILKCNNKTFS GTGPCQNVST
 96BW11_06 NCNTSTITQS CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
  96BW1210 NCNTSTITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCTNVST
 96BW15B03 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCKNNTFN GTGPCQNVST
 96BW16_26 NCNTSALTQA CPKVTFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
 96BW17A09 NCNTSTITQA CPKVSFDPIP IHYCAPAGYA ILECNNKTFN GTGPCTNVST
 96BWMO1_5 NCDTSTITQS CPKVTFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
 96BWMO3_2 NCNTSSITQA CPKVSFDPIP IHYCAPAGFA ILKCNNKTFN GTGPCQNVST
98BWMC12_2 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
98BWMC13_4 NCNTSTITQS CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GIGPCNNVST
98BWMC14_a NCNTSAISQA CPKISFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCQNVST
98BWMO14_1 NCNTSTVKQA CPKVTFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
98BWMO18_d NCNTSVITQA CPKVTFEPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
98BWMO36_a NCNTSTLTQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
98BWMO37_d NCNTSTVTQA CPKVSFDPIP IHYCTPAGYA ILKCNNKTFN GKGPCNNVST
99BW3932_1 NCYTSAIAQT CPKISFDPIP IHYCAPAGYA ILKCYNKTFN GTGPCKNVST
99BW4642_4 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCQNVST
99BW4745_8 NCNTSAVTQA CPKVSFDPIP IHYCAPAGYA ILKCNNETFN GTGPCNNVST
99BW4754_7 NCNTSAITQA CPKVTFDPIP IHYCAPAGYA ILKCNNETFN GTRPCNNVST
99BWMC16_8 NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNDVST
A2_CD_97CD SCNTSVITQA CPKVSFEPIP IHYCAPAGYA ILKCKDKEFN GKGSCSNVSS
A2_CY_94CY NCNTSTITQA CPKVTFEPIP IHYCAPAGFA ILKCKDPRFN GTGSCKNVSS
A2D___97KR HCDTSTITQA CPKVSFEPIP IHYCAPAGFA ILKCKDPEFN GTGPCKNVSS
A2G_CD_97C KCNTSTITQA CPKVSFEPIP IHYCAPAGFA ILKCKDPKFN GTGPCENVSS
A_BY_97BL0 NCNTSAMTQA CPXVTFEPIP IYYCAPAGFA ILKCXDTNFT RTXPCKNVST
A_KE_Q23_A NCNTSAITQA CPKVSFEPIP IHYCTPAGFA ILKCKDEGFN GTGLCKNVST
A_SE_SE659 NCNTSAITQA CPKVTFEPIP IHYCAPAGFA ILKCKDDAFN GTGRCKNVST
A_SE_SE725 NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCRDKEFN GTGPCNNVST
A_SE_SE753 NCNTSAITQA CPKVTFEPIP IHYCAPAGFA ILKCNDKEFN GTGLCNNVST
```

```
A_SE_SE853   NCNTSAITQA CPKVTFEPIP IHYCAPAGFA ILKCRDENFN GTGPCKNVST
A_SE_SE889   NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCNDKEFN GTGPCKNVSS
A_SE_UGSE8   NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCNEDEFN GTGTCRNVST
A_UG_92UG0   NCNTSALTQA RPKVTFEPIP IHYCAPAGYA ILKCNDKEFN GTGLCKNVST
A_UG_U455_   NCNTSTITQA CPKVSFEPIP IHYCAPAGFA ILKCKDPEFN GKGPCRNVST
AC_IN_2130   NCNTSAITQA CPKVTFNPIP IHYCTPAGYA ILKCNEETFN GTGPCKNVST
AC_RW_92RW   NCNTSAITQA CPKVSFEPIP INYCAPAGFA ILKCKDKKFN GTGPCKNVST
AC_SE_SE94   NCNTSAITQA CPKVSFDPIP IHYCAPPGFA ILKCKDAKFN GIGPCNNVST
ACD_SE_SE8   NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCNNKEFN GTGPCKNVST
ACG_BE_VI1   NCNTSAITQA CPKVSFDPIP IHYCAPAGFA ILKCNNKTFN GTGPCNNVST
AD_SE_SE69   NCNTSAIKQA CPKVTFEPIP IHYCAPAGFA ILKCKDTEFN GTGPCKNVST
AD_SE_SE71   NCNTSAITQA CPKVTFEPIP IHYCAPAGFA ILKCKDEKFN GTGPCKNVST
ADHK_NO_97   NCNTSVITQA CPKVTFEPIP IHYCAPAGFA ILKCNNKTFS GTGPCKNVST
ADK_CD_MAL   NCNTSVITQA CPKVTFDPIP IHYCAPAGFA ILKCNDKKFN GTEICKNVST
AG_BE_VI11   NCNSSTIKQA CPKVSFEPIP IHYCAPAGFA ILRCRDKKFN GTEPCKNVST
AG_NG_92NG   NCNVSTIKQA CPKVSFDPLP IHYCAPAGFA ILKCRGKNFT GTGQCKNVSS
AGHU_GA_VI   HCNTSTITQA CPKVSFDPIP IHYCAPAGFA ILKCKDKAFN GTGPCKNVST
AGU_CD_Z32   NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCRDEEFE GKGPCRNVST
AJ_BW_BW21   KCNTTVIKQA CPKVSFQPIP IHYCAPAGFA ILQCNDKKFN GTGPCKNVST
B_AU_VH_AF   NCNTSVITQA CPKVSFEPIP IHYCAPAGFA ILKCNNKTFN GKGPCANIST
B_CN_RL42_   NCNTSVITQA CPKVSFEPIP IHYCTPAGFA IIKCNNKKFN GTGPCTNVST
B_DE_D31_U   SCNTSVITQA CPKVSFEPIP IHYCTPAGFA MLKCKDKRFN GKGQCKNVST
B_DE_HAN_U   HCNRSVITQA CPKVSFEPIP IHYCAPAGFA ILKCNDKKFN GKGPCKNVST
B_FR_HXB2_   SCNTSVITQA CPKVSFEPIP IHYCAPAGFA ILKCNNKTFN GTGPCTNVST
B_GA_OYI__   HCNTSTITQA CPKISFEPIP MHYCTPAGFA ILKCNDKKFN GTGPCTNVST
B_GB_CAM1_   HCNTSVITQA CPKVSFEPIP IHYCTPAGFA ILKCNDKKFN GKGPCTNVST
B_GB_GB8_C   ECNASVITQA CPKISFEPIP IHFCAPAGFA ILKCNNKTFD GKGPCTNVST
B_GB_MANC_   SCNTSTITQA CPKVSFEPIP IHYCTPAGFA ILKCNNKKFD GKGQCTNVST
B_KR_WK_AF   HCNTSVITQA CPKVSFEPIP IHYCAPAGFA ILQCNDKKFN GTGPCSNVST
B_NL_3202A   SCNTSVITQA CPKVSFEPIP IHFCTPAGFA LLKCNDKKFN GTGPCKNVST
B_TW_TWCYS   SCNASVIKQA CPKVSFEPIP IHYCTPAGFA IIKCNNKTFN GTGTCTNVST
B_US_BC_L0   SCNTSVTTQA CPKISFEPIP IHYCAPAGFA ILKCKDKKFN GTGSCKNVST
B_US_DH123   SCNTSTLTQA CPKVSFEPIP IHYCAPAGFA ILKCKDKKFN GTGPCTNVST
B_US_JRCSF   SCNTSVITQA CPKVSFEPIP IHYCAPAGFA ILKCNNKTFN GKGQCKNVST
B_US_MNCG_   SCNTSVITQA CPKISFEPIP IHYCAPAGFA ILKCNDKKFS GKGSCKNVST
B_US_P896_   SCNTSVITQA CPKVSFQPIP IHYCVPAGFA MLKCNNKTFN GSGPCTNVST
B_US_RF_M1   HCNSSVITQA CPKVSFEPIP IHYCTPAGFA ILKCNDKKFN GTGPCKNVST
B_US_SF2_K   HCNRSVITQA CPKVSFEPIP IHYCTPAGFA ILKCNNKTFN GKGPCTNVST
B_US_WEAU1   NCKSSTITQA CPKVSFEPIP IHYCAPAGFA ILKCNDKKFN GKGPCKNVST
B_US_WR27_   HCNTSTITQA CPKVSFEPIP IHYCTPAGFA ILKCNDKKFN GTGQCKNVST
B_US_YU2_M   SCNTSVITQA CPKVSFEPIP IHYCAPAGFA ILKCNDKKFN GTGPCTNVST
BF1_BR_93B   NCNTSTLTQA CPKVSWDPIP IHYCAPAGYA ILKCNDKKFN GTGPCRNVST
C_BR_92BR0   NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
C_BW_96BW0   NCNTSTITQA CPKVTFDPIP IHYCAPAGYA ILKCNNKTFN GSGPCNNVST
C_BW_96BW1   NCNTSTITQS CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
C_BW_96BW1   NCNTSTITQA CPKVSFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCTNVST
C_BW_96BW1   NCNTSAITQA CPKVSFDPIP IHYCAPAGYA ILKCKNNTFN GTGPCQNVST
C_ET_ETH22   NCNTSTITQA CPKVSLDPIP IHYCAPAGYA ILKCRDKTFT GTGPCHNVST
C_IN_93IN1   NCNTSAITQA CPKVTFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCNNVST
C_IN_93IN9   NCNTSAITQA CPKVTFDPIP IHYCTPAGYA ILKCKDKTFN GTGPCHNVST
C_IN_93IN9   NCNTSAITQA CPKVTFDPIP IHYCTPAGYA ILKCNEKTFN GTGLCQNVST
C_IN_94IN1   SCNTSVITQA CPKVSFDPIP IHYCTPAGYA ILKCNDKTFN GTGPCRNVST
C_IN_95IN2   NCNTSALTQA CPKVTFDPIP IHYCAPAGYA ILKCNNKTFN GTGPCHNVST
CRF01_AE_C   HCNTSVIKQA CPKVTFDPIP IHYCTPAGYA ILKCNEKNFN GTGPCKNVSS
CRF01_AE_C   HCNTSVIKQA CPKVSFDPIP IHYCAPAGYA ILKCNDKNFN GTGPCKNVSS
CRF01_AE_C   HCNTSVIKQA CPKVSFDPIP IHYCAPAGYA ILKCNDKNFN GTGPCKKVSS
CRF01_AE_T   NCSTSVIKQA CPKISFDPIP IHYCTPAGYV ILKCNDKNFN GTGPCKNVSS
CRF01_AE_T   NCNTSVIKQA CPKISFDPIP IHYCTPAGYA ILKCNDKNFN GTGPCKNVSS
CRF01_AE_T   NCNTSVIKQA CPKVSFDPIP IHYCTPAGYA IIKCNDKNFN GTGPCKNVSS
CRF01_AE_T   NCNDSVIKQA CPKISFDPIP IHYCTPAGYA MLKCNDKNFN GTGPCKNVSS
CRF01_AE_T   NCNTSVIKQA CPKISFDPIP IHYCTPAGYA ILKCNDKNFN GTGPCKNVSS
```

```
CRF01_AE_T  NCNTSVIKQA CPKVSFDPIP IHYCTPAGYA ILKCNDKNFN GTGPCKNVSS
CRF02_AG_F  NCHTSAITKA CPRVTLEPIP IHYCAPAGLA ILKCNDKKFN GTGLCKNVST
CRF02_AG_F  NCNTSAIKQA CPKVTFEPIP IHYCAPAGFA ILKCNDKKFN GTGPCKNVST
CRF02_AG_G  NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCKDRNFN GTGPCQNVST
CRF02_AG_N  NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCKDKGFN GTGPCKNVST
CRF02_AG_S  NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCNDEKFN GTGPCKNVST
CRF02_AG_S  NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCNEKDFS GTGSCKNVST
CRF03_AB_R  SCNTSVVTQA CPKISFEPIP IHYCAPAGFA ILKCNDKKFN GTGPCTNVST
CRF03_AB_R  SCNTSVVTQA CPKISFEPIP IHYCAPAGFA ILKCNDKKFN GTGPCTNVST
CRF04_cpx_  NCNASTIKQA CPKVSFEPIP IHYCAPAGFA ILKCNEKNFT GLGPCTNVSS
CRF04_cpx_  SCNTSNHKQA CPKVTLEPNS HTLLCPGWFC DLKCNDKNSP GLGSCTNVSP
CRF04_cpx_  NCNVSSITQA CPKIKFEPIP IHYCAPAGFA ILQCNEKRFN GSGPCKNVSS
CRF05_DF_B  NCNTSTIKQA CPKVSWDPIP IHYCAPAGYA ILKCNEKKFS GTGPCKNVST
CRF05_DF_B  NCNTSTIKQA CPKVCWDPIP IHYCAPAGYA ILKCKEKRFN GTGPCKNVST
CRF06_cpx_  NCNVSTIKQA CPKVSFEPIP IHYCAPAGFA ILKCRDKDFN GTGPCKNVST
CRF06_cpx_  NCNASTIKQA CPKVTFEPIP IHYCAPAGFA ILKCRDKNFN GTGPCKNVST
CRF06_cpx_  NCNTSTIKQA CPKVSFEPIP IHYCAPAGFA ILKCRDKNFN GTGPCKNVST
CRF06_cpx_  NCNASTIKQA CPKVSFEPIP IHYCAPAGFA ILKCRDKNFN GTGACKNVST
CRF11_cpx_  NCNVSTIKQA CPKVTFEPIP IHYCAPAGFA ILKCRDKEFN GTGPCKNVST
CRF11_cpx_  NCNVSAVKQA CSKVTFEPIP IHYCAPAGFA ILKCRDKEFN GTGPCKNVST
D_CD_84ZR0  NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCNDKRFN GTGPCKNVSS
D_CD_ELI_K  NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCRDKKFN GTGPCTNVST
D_CD_NDK_M  NCDTSTITQA CPKISFEPIP IHFCAPAGFA ILKCRDKKFN GTGPCSNVST
D_UG_94UG1  NCNTSAITQA CPKMTFEPIP IHYCAPAGFA ILKCNEKKFN GTGPCKNVST
F1_BE_VI85  NCNTSTITQA CPKVSWDPIP IHYCAPAGYA ILKCNDKRFN GTGPCKNVST
F1_BR_93BR  NCDASTITQA CPKVSWDPIP IHYCAPAGYA ILKCNEKNFT GTGSCKNVST
F1_FI_FIN9  TCNTSTITQA CPKVSWDPIP IHYCAPAGYA ILKCKDKRFN GTGPCRNVST
F1_FR_MP41  NCNTSTIKQA CPKVSWDPIP IHYCAPAGYA ILKCRDPRFN GTGPCKNVST
F2_CM_MP25  NCNTSTLTQA CPKVSFDPIP IHYCAPAGFA ILKCNNKTFN GTGLCRNVST
F2KU_BE_VI  NCNTSAITQA CPKVSFEPIP IHYCAPAGFA ILKCNEKEFN GTGPCKNVST
G_BE_DRCBL  NCNVSTIKQA CPKVTFEPIP IHYCAPAGFA ILKCVDKKFN GTGTCNNVST
G_NG_92NG0  HCNVSTIKQA CPKVNFDPIP IHYCAPAGFA ILKCRDKEYN GTGPCKNVST
G_SE_SE616  HCNVSTIKQA CPKVTFDPIP IHYCAPTGFA ILKCRDKEFN GTGPCKNVST
H_BE_VI991  NCNTSVIKQA CPKVSFEPIP IHYCAPAGFA ILKCNGKKFN GTGPCTNVST
H_BE_VI997  NCNTSVITQA CPKVSFDPIP IHYCAPAGFA ILKCNNKTFN GTGPCTNVST
H_CF_90CF0  NCNTSVITQA CPKVSFEPIP IHYCAPAGFA ILKCNNKTFN GTGLCTNVST
J_SE_SE702  NCNTSVIKQA CPKVSFQPIP IHYCAPAGFA ILKCNNKTFN GTGPCKNVST
J_SE_SE788  NCNTSVIKQA CPKVSFQPIP IHYCAPAGFA ILKCNDKNFN GTGSCKNVST
K_CD_EQTB1  NCNTSTVTQA CPKVSFEPIP IHYCAPAGFA ILKCNNNTCN GTGPCTNVST
K_CM_MP535  NCNTSTITQT CPKVTFEPIP IHYCAPAGFA ILKCKDKRFN GTGPCKNVST
N_CM_YBF30  NCNTTAVTQA CPKTSFEPIP IHYCAPPGFA IMKCNEGNFS GNGSCTNVST
O_CM_ANT70  NCNSTTITQA CPKVSFEPIP IHYCAPAGYA IFKCNSTEFN GTGTCRNITV
O_CM_MVP51  NCNSTIIKQA CPKVSFEPIP IHYCAPTGYA IFKCNDTDFN GTGLCHNISV
O_SN_99SE_  NCNSTTIKQA CPKVSFEPIP IHYCAPAGYA IFKCNSTEFN GTGPCNNITA
O_SN_99SE_  NCNSTTIKQT CPKVSFEAIP IHYCAPAGYA IFKCNNTGFN GTGPCTNVTV
U_CD___83C  NCNTSTITQA CPKVSFEPIP IHYCAPAGFA ILKCNDKKFN GTGPCKNVST
```

```
           301                                                        350
00BW0762_1 VQCTHGIKPV VSTQLLLNGS LSEE.GIVIR SENITNNAKT IIVHLNESVE
00BW0768_2 VQCTHGIKPV VPTQLLLNGS LAEE.EIIIR SEALTNNAKT IIARLNKSVE
00BW0874_2 VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLAKNDKT IIVHLNESVE
00BW1471_2 VLCTHGIKPV VSTQLLLNGS LSEG.EIMIR SENLTDNAKT IIVQLNKPVE
00BW1616_2 VQCTHGIKPV VSTQLLLNGS LAEE.EIVIR SENITNNAKI IIVHLNETVD
00BW1686_8 VQCTHGIKPV VSTQLLLNGS LAEG.EIIIR SENMTDNVKP IIVHLNESVE
00BW1759_3 VQCTHEIKPV VSTQLLLNGS LAEE.QIIIR SENMTNNAKT IIVHLKKPVQ
00BW1773_2 VQCTHGIKPV VSTQLLLNGS LAEK.EIIIR SENLTNNAKI IIVHLKDPVN
00BW1783_5 VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTNNVKT IVVHLNESIQ
00BW1795_6 VQCTHGIKPV VSTQLLLNGS LAEG.EIIIR SKNLTDNART IIVHLNESVQ
00BW1811_3 VQCTHGIKPV VSTQLLLNGS LAEE.DIIIR SENLTDNVKT IIVHLNESVE
00BW1859_5 VQCPHGIKPV VSTQLLLNGS LAEE.EIIIR SENIEDNVKT IIVHLNESIE
00BW1880_2 VQCTHGIKPV VATQLLLNGS LAEE.EIVIR FKNITNNAKI IIVQLNTSVG
00BW1921_1 VQCTHGIKPV VSTQLLLNGS LAEE.GIIIR SENLTDNAKT ITVQLDQAVE
00BW2036_1 VQCTHGIRPV VSTQLLLNGS LAEG.EIIIR SENLADNAKT IIVHFNESVE
00BW2063_6 VQCTHGIKPV VSTQLLLNGS LAEK.EIIIR SKNITDNVKT IIVHLNEAVE
00BW2087_2 VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTNNAKT IIVHLNDSVE
00BW2127_2 VQCTHGIKPL VSTQLLLNGS LAQQ.QIMIT SENLTNNAKI IIVHLKEAIN
00BW2128_3 VQCTHGIKPV VSTQLLLNGS LAEE.IIIR. SENLTNNAKT IIVQLKEPVK
00BW2276_7 VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTNNVKT IIVHLNKSVE
00BW3819_3 VQCTHGIKPV VSTQLLLNGG LAEK.EIIIR SENLTNNVKT IIVHLNESVE
00BW3842_8 VQCTHGIKPV VSTQLLLNGS LAEG.EIIIR SKNLSDNAKI IIVHLNESVG
00BW3871_3 VQCTHGIKPV VSTQLLLNGS IAEG.EIIIR SENLTNNAKT IIVHLNESVK
00BW3876_9 IQCTHGIKPV VSTQLLLNGS LAEK.DIIIR SENLTNNIKT IIVHLNDSVQ
00BW3886_8 VQCTHGIKPV VSTQLLLNGS LAEG.EIIIR SENLTNNAKT IIVHLNESIE
00BW3891_6 VQCTHGIKPV VSTQLLSNGS LSEK.EIIIR SENLTNNVKT IIVHLNESVE
00BW3970_2 IQCTHGIKPV VSTQLLLNGS RAEK.EIIIR SENMTNNAKT IIVHLNESIE
00BW5031_1 VQCTHGIKPV VSTQLLLNGS LA.E.EIVIR SENISDNAKT IIVHLNESVE
 96BW01B21 VQCTHGIKPV VSTQLLLNGS LAEE.DIIIR SENLTDNVKT IIVHLNESIE
  96BW0407 VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTDNAKI IIVHLNESVR
  96BW0502 VQCAHGIKPV VSTQLLLNGS VAKG.EIIIR SENLTNNAKI IIVQLNKPVK
 96BW06_J4 VQCTHGIKPV VSTQLLLNGS LAEE.KIIIR SGNVTNNAKT IIVHLNESVE
 96BW11_06 VQCTHGIKPV VSTQLLLNGS LAEK.EIIIR FKNITNNAKT IIVHLNESVE
  96BW1210 VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENITDNVKT IIVHLNKSVE
 96BW15B03 VQCTHGIKPV VSTQLLLNGS LAEG.EIIIR SKNLSNNAYT IIVHLNDSVE
 96BW16_26 VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTNNAKT LIVHLNESVE
 96BW17A09 VQCTHGIKPV VSTRLLLNGS LAEE.EIIIR SENLTNNAKI IMVQLNESIE
 96BWMO1_5 VQCTHGIKPV VSTQLLLNGS LAEK.EIIIR SKNITDNVKT IIVHLNESVE
 96BWMO3_2 VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTDNAKT IIVHFNESVQ
98BWMC12_2 VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTNNVKT IIVHLNESVE
98BWMC13_4 VQCTHGIKPV VSTQLLLNGS LAEK.EIVIR SENLTDNVKT IIVHLNQSVQ
98BWMC14_a VQCTHGIKPV VSTQLLLNGS LAEE.EVVIR SENLTNNAKT IIVQLKEPVK
98BWMO14_1 VQCTHGIKPV VSTQLLLNGS LAEK.EIIIR SENITNNIKT IIVHLNEPVE
98BWMO18_d VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SKNITNNANT IIVQLKDPVD
98BWMO36_a IQCTHGIKPV VSTQLLLNGS IAEE.EIIIR SENLTDNVKT IIVHLKDPID
98BWMO37_d VQCTHGIKPV VSTQLLLNGS IAEE.EIIIR SENLTNHFRT IIVQLNQSIE
99BW3932_1 VQCTHGIKPV VSTQLLLNGS LAEE.EIVIR SENLTNNVKT IIVHLNESIE
99BW4642_4 VQCTHGIKPV VSTQLLLNGS LAEG.EIIVR SENLTDNVKT IIVHLDKPVG
99BW4745_8 VQCTHGIKPV VSTQLLLNGS LAEK.EIIVR SENLTNNAKI IIVQLKQSVG
99BW4754_7 VQCTHGIKPV VSTQLLLNGS RAEK.EVVIR AESLTDNAKV IIVHLKDSVQ
99BWMC16_8 VQCTHGIKPV VSTQLLLNGS LAEE.DIIIR SKNLTDNAKT IIVHLNESVN
A2_CD_97CD VQCAHGIRPV ASTQLLLNGS LAEG.KVMIR SENITDNAKN IIVQFNKPVP
A2_CY_94CY VQCTHGIKPV ASTQLLLNGS LAEGGKIMIR SENITNNAKN IIVQFTKPVL
A2D___97KR VQCTHGIRPV ASTQLLLNGS LAEGNKTIIR SANITDNTKN IIVQFTKPVQ
A2G_CD_97C VQCTHGIRPV VSTQLLLNGS LAEE.EVMIR SENITDNAKN IIVQFDKPVE
A_BY_97BL0 VQCTHXIKPV VSTQLLLNGS LAEKX.VMIR SENITDNVKI IIVQLTEPVN
A_KE_Q23_A VQCTHGIKPV VSTQLLLNGS LAEKN.ITIR SENITNNAKI IIVQLVQPVT
A_SE_SE659 VQCTHGIKPV VSTQLLLNGS LAKGG.IRIR SENITNNVKT ILVQLDKPVN
A_SE_SE725 VQCTHGIKPV VSTQLLLNGS LAEEK.IMIR SENISDNAKT IIVQLTEPVT
A_SE_SE753 VQCTHGIKPV VSTQLLLNGS LATK..IMIR SENITNNAKT IIVQLVEPVE
```

```
A_SE_SE853   VQCTHGIKPV VSTQLLLNGS LAREK.VMIR SENITNNVKN IIVQLKEPVE
A_SE_SE889   VQCTHGIRPV VSTQLLLNGS LAETE.VMIR AENITNNIKN IIVQFNKSVE
A_SE_UGSE8   VQCTHGIRPV VSTQLLLNGS LAKEE.VRIR SENISDNAKT IIVQFTKPVE
A_UG_92UG0   VQCTHGIRPV VSTQLLLNGS LAEGK.VMIR SENITNNVKN IIVQLNESVT
A_UG_U455_   VQCTHGIKPV VSTQLLLNGS LAERE.IRIR SENFTNNAKT IIVQLVNPVK
AC_IN_2130   VQCTHGIKPV VSTQLLLNGS LAEKE.VIIR SENITNNVKN IIVQLAEPVR
AC_RW_92RW   VQCTHGIKPV VSTQLLLNGS LAEEE.IIIR SENITNNAKT IIVQLNETVQ
AC_SE_SE94   VQCTHGIKPV ISTQLLLNGS LSETG.VKIR SENITNNAKT IIVQLDEAVE
ACD_SE_SE8   VQCTHGIKPV VSTQLLLNGS LAEEE.IIVR SENLTNNAKI IIIQLNETVK
ACG_BE_VI1   VQCTHGIKPV VSTQLLLNGS LAEK..VVIR SENITNNAKT IIVQFDSPVK
AD_SE_SE69   VQCTHGIKPV VSTQLLLNGS LA.EGKVRIR SENITDNTKN IIVQFTEPVT
AD_SE_SE71   VQCTHGIKPV VSTQLLLNGS LAKEE.VIIR SENITNNAKN IIVQFVKPVT
ADHK_NO_97   VQCTHGIKPV VSTQLLLNGS LA..EKVIIR SKNITDNTKN IIVHFNESVQ
ADK_CD_MAL   VQCTHGIKPV VSTQLLLNGS LA.EEEIMIR SENLTDNTKN IIVQLNETVT
AG_BE_VI11   VQCTHGIKPV VSTQLLLNGS LA.EEEIIVR SENFTNNAKV IIVQLKEPIE
AG_NG_92NG   VQCTHGIKPV VSTQLLLNGS LA.EGEIVIR SENLTDNAKV IIVQLNKTIG
AGHU_GA_VI   VQCTHGIRPV VTTQLLLNGS LA.EGEIIIR SENITENTKN IIVQLNETVE
AGU_CD_Z32   VQCTHGIKPV VSTQLLLNGS LAEKE.VRIR SENFSDNAKI IIVQLAKPVN
AJ_BW_BW21   VQCTHGIKPV VSTQLLLNGS IA.EEEIIIR SENITNNAKT IIVQLNNTVE
B_AU_VH_AF   VQCTHGIRPV VSTQLLLNGS LA.EKEIVIR SDNFTDNAKS IIVQLNESVE
B_CN_RL42_   VQCTHGIRPV VSTQLLLNGS LA.EEEVVIK FSNFTDNARV IIVQLNESVE
B_DE_D31_U   VQCTHGIRPV VSTQLLLNGS LA.EEEVVIR SDNFTDNAKT IIVQLKESVE
B_DE_HAN_U   VQCTHGIRPV VSTQLLLNGS LA.EKEVVIR SDNFTDNTKT IIVHLNESVE
B_FR_HXB2_   VQCTHGIRPV VSTQLLLNGS LA.EEEVVIR SVNFTDNAKT IIVQLNTSVE
B_GA_OYI__   VQCTHGIKPV VSTQLLLNGS LA.EEEVIIR SSNFTNNAKI IIVQLNKSVE
B_GB_CAM1_   VQCTHGIRPV VSTQLLLNGS LA.EKEVVIR SENFTNNAKT IIVQLKEPVE
B_GB_GB8_C   VQCTHGIRPV VSTQLLLNGS LA.EEKVVIR SDNFTDNVKT IIVQLKEAVE
B_GB_MANC_   IQCTHGIRPV VSTQLLLNGS LA.EEEVVLR SDNFTDNAKT IIVHLNESVE
B_KR_WK_AF   VQCTHGIRPV VSTQLLLNGS LA.EEEIVLR SENFTNNAKT IIVQLNASVE
B_NL_3202A   VQCTHGIRPV VSTQLLLNGS LA.EEEVVIR SANFSNNAKT IIVQLNESVA
B_TW_TWCYS   VQCTHGIRPV VSTQLLLNGS IA.EEEILIK SENITNNAKT IIIQLNKSVK
B_US_BC_L0   VQCTHGIKPV VSTQLLLNGS LA.EEEVVIR SANFSDNAKT IIVQLKEAVE
B_US_DH123   VQCTHGIRPV VSTQLLLNGS LA.EEEVVIR SSNFTDNAKI IIVQLNETVE
B_US_JRCSF   VQCTHGIRPV VSTQLLLNGS LA.EEKVVIR SDNFTDNAKT IIVQLNESVK
B_US_MNCG_   VQCTHGIRPV VSTQLLLNGS LA.EEEVVIR SENFTDNAKT IIVHLNESVQ
B_US_P896_   VQCTHGIRPV VSTQLLLNGS LA.EEDIVIR SENFTDNAKT IIVQLNESVV
B_US_RF_M1   VQCTHGIRPV VSTQLLLNGS LA.EEEVVIR SENFTDNVKT IIVQLNASVQ
B_US_SF2_K   VQCTHGIRPI VSTQLLLNGS LA.EEEVVIR SDNFTNNAKT IIVQLNESVA
B_US_WEAU1   VQCTHGIRPV VSTQLLLNGS LA.EEDIVIR SENFTDNAKN IIVQLNVSIE
B_US_WR27_   VQCTHGIRPV VSTQLLLNGS LA.EEEVVIR SANFTNNAKT IIVQLKESVE
B_US_YU2_M   VQCTHGIRPV VSTQLLLNGS LA.EEEIVIR SENFTNNAKT IIVQLNESVV
BF1_BR_93B   VQCTHGIKPV VSTQLLLNGS LA.EKDIIIR SQNISDNAKT IIVQLNVSVP
C_BR_92BR0   IQCTHGTKPV VSTQLLLNGS LAEE.EIIIR SKNLTDNVKT IIVHLNESVE
C_BW_96BW0   VQCTHGIKPV VSTQLLLSGS LAEE.EIVIR SENLTNNAKI IIVHLNKTVR
C_BW_96BW1   VQCIHGIKPV VSTPLLLNGS LAEK.EIIIR FKIITNNAKT IIVHLNESVE
C_BW_96BW1   VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENITDNVKT IIVHLNKSVE
C_BW_96BW1   VQCTHGIKPV VSTQLLLNGS LAEGGEIIIR SKNLSNNAYT IIVHLNDSVE
C_ET_ETH22   VQCTHGIKPV VSTQLLLNGS IAEG.ETIIR FENLTNNAKI IIVQLNESVE
C_IN_93IN1   VQCTHGIKPV VSTQLLLNGS LAEG.EIIIR SENLTNNVKT IIVHLNQSVE
C_IN_93IN9   VQCTHGIKPV VSTQLLLNGS LAEG.EIIIR SENLTNNAKT IIVHLNQSVQ
C_IN_93IN9   VQCTHGIKPV VSTQLLLNGS LAEG.EIIIR SENLTDNVKT IIVHLNQSVE
C_IN_94IN1   VQCTHGIKPV VSTQLLLNGS LSEG.EIIIR SENLTNNVKT IIVHLNKSVE
C_IN_95IN2   VQCTHGIKPV VSTQLLLNGS LAEG.GIIIR SENLTNNVKT IIVHLNQPVE
CRF01_AE_C   VQCTHGIKPV VSTQLLLNGS LAEE.DIIIR SENLTNNAKT IIVHLNKSVE
CRF01_AE_C   VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SEDLTDNAKT IIVHLNKSIE
CRF01_AE_C   VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTNNAKT IIVHLNKSVE
CRF01_AE_T   VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTNNAKT IIVHLNKSVE
CRF01_AE_T   VQCTHEIKPV VSTQLLLNGS LAEE.EIIIR FENLTNNAKT IIVHLNKSVE
CRF01_AE_T   VQCTHGIKPV VSTQLLLNGS LAEE.KIIIR SENLTNNAKT IIVHLHESVE
CRF01_AE_T   VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SENLTNNAKT IIVHLNKSVE
CRF01_AE_T   VQCTHGIKPV VSTQLLLNGS LAEE.EIIIR SEDLTNNAKT IIVHLNKSVE
```

```
CRF01_AE_T  VQCTHGIKPV VSTQLLLNGS LAKE.EIIIR SENLTNNAKT IIVHLNKSVK
CRF02_AG_F  VQCTHGIKPV VSTQLLLNGS LAEEE.VVIR SENITNNAKN IIVQLVAPVR
CRF02_AG_F  VQCTHGIKPV VSTQLLLNGS LAEEE.VVIR SENITNNAKN IIVQLVTPVR
CRF02_AG_G  VQCTHGIKPV VSTQLLLNGS LAEEE.IVIR SENITNNVKN IIVQLAKPVR
CRF02_AG_N  VQCTHGIKPV VSTQLLLNGS LAEGE.VVIR SENITNNAKT IIVQLANPVK
CRF02_AG_S  VQCTHGIKPV VSTQLLLNGS LAEEE.IVIR SENFTNNAKI IIVQLHESVK
CRF02_AG_S  VQCTHGIKPV VSTQLLLNGS LAEGD.IVIR SENISNNAKT IIVQLNKPVW
CRF03_AB_R  VQCTHGIKPV VSTQLLLNGS LA.EEEVVIR SVNFTDNTKT IIVQLKEPVE
CRF03_AB_R  VQCTHGIKPV VSTQLLLNGS LA.EEEVVIR SVNFTDNTKT IIVQLKEPVE
CRF04_cpx_  VRCTHGIKPV VSTQLLLNGS LA.TEEVVIR SKNITDNTKN IIVQLAKAVK
CRF04_cpx_  VQCTHGIKPV VSTQLLLNGS LA.TGGVVIR SKNFTDNPKN IIVQLDKAVK
CRF04_cpx_  VQCTHGIKPV VSTQLLLNGS LS.TEGVVLR SKNFTDNTKN IIVQLAEAVK
CRF05_DF_B  VQCTHGIKPV VSTQLLLNGS LA.KEGIIIR SQNISDNAKN IIVHLNESVH
CRF05_DF_B  VQCTHGIKPV VSTQLLLNGS LA.EESIIIR SQNILDNTKT IIVHLNESVQ
CRF06_cpx_  VQCTHGIKPV VSTQLLLNGS LA.EGNITIK TENITDNTKN IIVQLNQPVE
CRF06_cpx_  VQCTHGIKPV VSTQLLLNGS LA.EEEIIIK SKNLTDNTKI IIVQLNKSVE
CRF06_cpx_  VQCTHGIKPV VSTQLLLNGS LA.EDEIIIK SENHTNNAKI IIVQLNKTVQ
CRF06_cpx_  VQCTHGIKPV VSTQLLLNGS LA.EEEIIIK TENLTDNSKN IIVQLNKSIE
CRF11_cpx_  VQCTHGIKPV VSTQLLLNGS LA.EEKVKIR SENFTNNAKT IIVQFNNTVR
CRF11_cpx_  VQCTHGIKPV VSTQLLLNGS LA.EGEVRIR SENLTNNAKT IIVQLNSTVR
D_CD_84ZR0  VQCTHGIRPV VSTQLLLNGS LA.EEEIVIR SENLTNNAKI IIVHLNQSVE
D_CD_ELI_K  VQCTHGIRPV VSTQLLLNGS LA.EEEVIIR SENLTNNAKN IIAHLNESVK
D_CD_NDK_M  VQCTHGIRPV VSTQLLLNGS LA.EEEIIIR SENLTNNVKT IIVQLNASIV
D_UG_94UG1  VQCTHGIKPV VSTQLLLNGS LA.EEEIIIR SENLTNNAKI IIVQLNESVP
F1_BE_VI85  VQCTHGIKPV VSTQLLLNGS LA.EEGIVIR SQNISNNAKT IIVHLNESVQ
F1_BR_93BR  VQCTHGIKPV VSTQLLLNGS LA.EGEIVIR SQNISDNAKT IIVHLNESVQ
F1_FI_FIN9  VQCTHGIRPV VSTQLLLNGS LS.EGGIIIR SQNLSDNAKT IIVHLNESVQ
F1_FR_MP41  VQCTHGIRPV VSTQLLLNGS LA.EEDIIIR SQNISDNAKT IIVHLNESVQ
F2_CM_MP25  VQCTHGIKPV VSTQLLLNGS LA.EEKMIIR SENISDNTKT IIVQFKNPVK
F2KU_BE_VI  VQCTHGIRPV ISTQLLLNGS LA.EKEIIIR SGNISDNTKN IIVQLNETVE
G_BE_DRCBL  VQCTHGIKPV VSTQLLLNGS LA.EKDIIIS SENISDNAKV IIVHLNRSVE
G_NG_92NG0  VQCTHGIKPV VSTQLLLNGS LA.EEDIRIR SENFTDNTKV IIVQLNNSIE
G_SE_SE616  VQCTHGIKPV VSTQLLLNGS LA.EGKIKVR SENFTDNTKV IIVQLNKTVE
H_BE_VI991  VQCTHGIRPV VSTQLLLNGS LAEVEEVIIR SKNITDNTKN IIVQLNEPVQ
H_BE_VI997  VQCTHGIKPV VSTQLLLNGS LA.EGQVIIR SKNISDNTKN IIVQLDSPIE
H_CF_90CF0  VQCTHGIRPV VSTQLLLNGS LA.EEQIIIR TKNISDNTKN IIVQLKTPVN
J_SE_SE702  VQCTHGIKPV VSTQLLLNGS VA.EGDIIIR SENISDNAKN IIVQLNDTVE
J_SE_SE788  VQCTHGIKPV VSTQLLLNGS IA.EGDIIIR SENISDNAKN IIVQLNKTVE
K_CD_EQTB1  VQCTHGIKPV VSTQLLLNGS LA.EEEIIIR SEDITKNTKN IIVQLNEAVE
K_CM_MP535  VQCTHGIKPV VSTQLLLNGS LA.EEEIIIR SENITDNTKN IIVQLNETVQ
N_CM_YBF30  VQCTHGIKPV ISTQLILNGS LNTDGIVIR. ....NDSHSN LLVQWNETVP
O_CM_ANT70  VTCTHGIRPT VSTQLILNGT LS.KGKIRMM AKDILEGGKN IIVTLNSTLN
O_CM_MVP51  VTCTHGIKPT VSTQLILNGT LS.REKIRIM GKNITESAKN IIVTLNTPIN
O_SN_99SE_  VTCTHGIKPT VSTQLILNGT LS.EGNIRIM GKNISDNMKN IIVTLNSTIN
O_SN_99SE_  VTCTHGIRPT VSTQLILNGT IS.EGKIRIM GKNISDTGKN IIVTINSTIN
U_CD___83C  VQCTHGIRPV VSTQLLLNGS LSEE.EVIIR SENITNNAKT IIVQLNETVK
```

```
            351                                                        400
00BW0762_1  IVCTRPGNN. .TRRSVRIG. .PGQTFYATG .....DIIGN IRQAHCNISK
00BW0768_2  IECTRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGD IREAHCNINK
00BW0874_2  IVCTRPFNN. .TRKSIRIG. .PGQTFYATG .....DVIGD IRQAHCNTSE
00BW1471_2  IVCTRPGNN. .TRKSVRIG. .PGQTFYATG .....DIIGD IRQAHCNISN
00BW1616_2  IVCTRPNNN. .TRKSMRIG. .PGQTFYATG .....DIIGN IREAHCNISK
00BW1686_8  INCTRPNNN. .TRKSIRIG. .PGQTFYATG .....AIIGD IRQAYCTVNA
00BW1759_3  IVCTRPNNN. .TRRSVRIG. .PGQTFYARG .....DIIGN IRQAHCNISR
00BW1773_2  ITCTRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGD IRQAHCNISA
00BW1783_5  INCTRPGNN. .TRKSVRIG. .PGQTFYATG .....DIIGD IRQAHCNISE
00BW1795_6  IVCTRPNNN. .TRKSVRLG. .PGQAFYATG .....DIIGD IRKAYCTINE
00BW1811_3  IVCTRPGNN. .TRKSIRIG. .PGQAFFATG .....EIIGD IRQAHCNISA
00BW1859_5  IVCIRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGN IREAHCNITR
00BW1880_2  IVCTRPNNN. .TKKSMRIG. .PGQTFYATG .....DIIGD IRQAHCNISE
00BW1921_1  IECTRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGN TRQAHCNVSA
00BW2036_1  IECIRPNNN. .TRKSIRIG. .PGQVFYATG .....DIIGD IREAHCNITE
00BW2063_6  IVCTRPGNN. .TRKSVRIG. .PGQTFYATG .....EVIGD IREAHCNISE
00BW2087_2  IVCTRPNNN. .TRKSIRIG. .PGQAFYATD .....AIIGD IRQAHCNISR
00BW2127_2  IVCTRPNNN. .TRTSIRIG. .PGHSFFATN .....GIIGD IRQAHCSISK
00BW2128_3  INCTRPNNN. .TRKSIRIG. .PGQAFYATG .....DIIGD IRQAHCNISK
00BW2276_7  IVCVRPNNN. .TRKSVRIG. .PGQTFFAT. .....NIIGD IREAHCNISE
00BW3819_3  IKCTRPNNN. .TRRSVRIG. .PGQAFYTN. .....DIIGD IRLAHCNISK
00BW3842_8  IVCTRPNNN. .TRKSIRIG. .PGQTFYAAG .....DIIGN IRQAHCNISE
00BW3871_3  ITCTRPNNN. .TRESIRIG. .PGQTFYATG .....DIIGD IRKAYCNISI
00BW3876_9  IVCTRPNNN. .IRKSVRIG. .PGQAFYATG .....DIIGD IREAYCNING
00BW3886_8  IVCVRPNNN. .TRKSIRIG. .PGQTFYATG .....EIIGN IRQAYCSISG
00BW3891_6  IECTRPNNN. .TRRSIRIG. .PGQTFYATG .....EIIGD IRQAYCTINE
00BW3970_2  IECIRPNNN. .TRKSIRIG. .PGQTFYATN .....GMIGD IRQAHCNISG
00BW5031_1  IECRRPNNN. .TGKSVRIG. .PGQTFFATG .....GIIGE IRRAHCDING
 96BW01B21  INCTRPNNN. .TRKSIRIG. .PGQTFYAAG .....EIIGK IRLAYCNISE
  96BW0407  IECTGPNNN. .TRKSMRIG. .PGQTFYATG .....EIVGD IRQAHCNISE
  96BW0502  IVCVRPNNN. .TRKSVRIG. .PGQTFYATG .....EIIGD IRQAYCIINK
 96BW06_J4  IVCTRPNNN. .TRKSIRIG. .PGQTFYAT. .....DIIGD IRQAYCNVSK
 96BW11_06  IVCIRPNNN. .TRKSVRIG. .PGQTFYATE .....AIIGN IREAHCNISE
  96BW1210  IVCTRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGD IRQAHCNISK
 96BW15B03  IVCTRPNNN. .TRKGIRIG. .PGQTFYATE .....NIIGD IRQAHCNISA
 96BW16_26  IVCIRPNNN. .TRKSIRIG. .PGQTFFATG .....DIIGD IRQAHCIING
 96BW17A09  IVCTRPNNN. .TRKSTRIG. .RGQTFYAMG .....RIIGD IRQAHCNISG
 96BWMO1_5  IECTRPGNN. .TRRSVRIG. .PGQAFYATG .....DIIGD IRAAHCNISE
 96BWMO3_2  INCTRPGNN. .TRKSVRIG. .PGQAFYATG .....DIIGD IRKAYCNISK
98BWMC12_2  IVCTRPNNN. .TRKSMRIG. .PGQIFYATG .....DIIGD IREAHCNISK
98BWMC13_4  IECTRPGNN. .TRKSMRIG. .PGQAFYATG .....EIIGN IRQAYCNINE
98BWMC14_a  IVCTRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGD IRQAHCNISE
98BWMO14_1  IVCTRPGNN. .TRTSIRIG. .PGQTFYATG .....DIIGD IRQAHCNISE
98BWMO18_d  ILCVRPSNN. .TRKSVRIG. .PGQTFYATG .....DIIGD IRQAHCNISA
98BWMO36_a  IVCTRPGNN. .TRKSVRIG. .PGQTFFATG .....DIIGD IRQAHCNISK
98BWMO37_d  INCTRPSNN. .TRKSIRIG. .PGQAFYATN .....DIIGD IRQAHCNISE
99BW3932_1  IVCIRPNNN. .TRKSIRIG. .PGQTFYATG .....AIIGN IREAYCNISG
99BW4642_4  IVCIRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGN IKEAYCNIKE
99BW4745_8  IECIRPNNN. .TRKSIRIG. .PGQTFYATG .....EIIGD IRKAHCTINK
99BW4754_7  INCTRPNNN. .TRKSMRIG. .PGQTFYATG .....EIIGD IRQAHCNISR
99BWMC16_8  ITCTRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGD IRQAHCSINK
A2_CD_97CD  INCTRPNNN. .TRKSIRFG. .PGQAFYTNN .....NIIGD IRQAHCNISI
A2_CY_94CY  ITCIRPNNN. .TRKSIRFG. .PGQAFYTN. .....EIIGD IRQAHCNINK
A2D___97KR  INCTRPDVG. .QRRSVRIG. .PGRAFYTRQ TYTR.QAKGD IRQAQCNISS
A2G_CD_97C  ITCIRPNNN. .TRKSIRFG. .PGQAFYTN. .....SIIGD IRQAYCNISK
A_BY_97BL0  ITCIRPGNN. .TRTSIRIG. .PGQTFYATG .....DVIXD IRKAYCNVSR
A_KE_Q23_A  IKCIRPNNN. .TRKSIRIG. .PGQAFYATG .....DIIGD IRQAHCNVTR
A_SE_SE659  ITCIRPYHN. .TRTRIHIG. .PG.RSFYTG .....DIKGS IRQAHCTVNR
A_SE_SE725  INCTRPSNN. .TRTSIRIG. .PGQAFYATG .....DITGD IRQAHCNVSR
A_SE_SE753  INCTRPNNN. .TRTSVPIG. .PGKVFYATG .....EIIGD IRQAHCNVSK
```

```
A_SE_SE853   INCTRPGNN. .TRKSIRIG. .PGQAFYATG .....EVIGD IRQAHCNVSR
A_SE_SE889   IICIRPNNN. .TRKSIRIG. .PGQAFYATG .....DIIGD IRQAYCDVNR
A_SE_UGSE8   IICTRPNNN. .TRKSIRIG. .PGQAFYGMG .....DIIGD IRKAHCNVSR
A_UG_92UG0   INCTRPNNN. .TRRSVRIG. .PGQTFYATG .....DIIGD IRQAHCNVSG
A_UG_U455_   INCSRPYNTR KNIRRYSIG. .SGQAFYVTG .....KIIGD IRQAHCNVSR
AC_IN_2130   INCTRPNNN. .TRTSIRIG. .PGQTFYTS. .....NIIGD IRQAHCNVSR
AC_RW_92RW   INCSRPNNN. .TRKSVHIG. .PGQAFYATG .....DVIGD IRQAYCTVNG
AC_SE_SE94   INCTRPGNN. .TRRSVHIG. .PGQAFYATG .....DITGD IRKAHCIVNG
ACD_SE_SE8   INCTRPNNN. .TRNSIRIG. .PGQAFYATG .....AITGD IRQAHCNVSR
ACG_BE_VI1   INCTRPGNN. .TRKSVRIG. .PGQTFYATG .....DIIGD IRQAHCNISG
AD_SE_SE69   INCTRPNNNT .RK.SVRIG. .PGQALYVTG GII..G...D IRQAFCEVNR
AD_SE_SE71   INCTRPNNN. .TRKSVHMG. .PGKVFYATG .....DIIGD IRQAHCNVSK
ADHK_NO_97   INCTRIANNT RKS..IHIG. .PGQAFYAAE PVI..G...D IRQAHCNISE
ADK_CD_MAL   INCTRPGNNT .RR.GIHFG. .PGQALYTTG IVG......D IRRAYCTINE
AG_BE_VI11   INCTRPNNNT RKSIGLGPG. ...QAFYATG DII..GD... IRQAHCNVSG
AG_NG_92NG   INCTRPNNNT RKSIRIGPG. ...QAFYATG EII..G.... ..........
AGHU_GA_VI   INCTRPNNNT RKG..IRIG. .PGRVIYATS AIT..G...D IRQAHCNISK
AGU_CD_Z32   ITCMRPNNY. .TRKSIHIG. .PGRALYPEG .....DIIGD IRQAHCNVSR
AJ_BW_BW21   IKCVRPANNT RKGIHTGPG. ...QVLYATG AVV..GD... IRQAHCNVSR
B_AU_VH_AF   IHCMRPNNNT .RK.GIYVG. .PGRHIYATE KIV..G...D IRQAHCNISR
B_CN_RL42_   IKCIRPNNNT .RK.SIHLG. .PGKAWYTTG QII..G...D IRQAHCNLSS
B_DE_D31_U   INCTRPNNYT .SK.RIRIG. .ARRAFYTKG KII..G...D IRQAHCNISG
B_DE_HAN_U   INCTRPNNNT .RK.GIHIG. .PGRAVYTTG RIV..G...D IRLAHCNISR
B_FR_HXB2_   INCTRPNNNT .RK.RIRIQR GPGRAFVTIG KIG......N MRQAHCNISR
B_GA_OYI__   INCTRPNNNT .RN.RISIG. .PGRAFHTTK QII..G...D IRQAHCNLSR
B_GB_CAM1_   INCTRLNNNT .RK.SIAIG. .PGRTVYATD RII..G...D IRQAHCNLSS
B_GB_GB8_C   INCTRPNNNT .RK.GIYMG. .PGRRFYTTG RII..G...D IRQAHCNISK
B_GB_MANC_   INCTRPSNNS .RK.SIYIG. .PGRRFHVTR AVT..G...D IRQAHCNISK
B_KR_WK_AF   INCTRLNNNT .RK.SIRIG. .PGSTFYATG AII..G...D IRQAHCNISR
B_NL_3202A   INCTRPNNNT .RK.GIHIG. .PGKAFYATG QII..G...D IRQAHCNLSR
B_TW_TWCYS   INCTRPNNIS KRR.SMHIG. .TGRVFYTQT ..I..G...N IRQAHCNLSK
B_US_BC_L0   INCTRPNKKT .RK.RITTG. .PGRVYYTTG EIV..G...D IRQAHCNLSR
B_US_DH123   INCTRPNNNT .RK.GITLG. .PGRVFYTTG EIV..G...D IRKAHCNISK
B_US_JRCSF   INCTRPSNNT .RK.SIHIG. .PGRAFYTTG EII..G...D IRQAHCNISR
B_US_MNCG_   INCTRPNYNK .RK.RIHIG. .PGRAFYTTK NII..G...T IRQAHCNISR
B_US_P896_   INCTRPNNNT .RR.RLSIG. .PGRAFYARR NII..G...D IRQAHCNISR
B_US_RF_M1   INCTRPNNNT .RK.SITKG. .PGRVIYATG QII..G...D IRKAHCNLSR
B_US_SF2_K   INCTRPNNNT .RK.SIYIG. .PGRAFHTTG RII..G...D IRKAHCNISR
B_US_WEAU1   INCTRPNNNT .RK.KITLG. .PGRVLYTTG EII..G...D IRRAHCNLSR
B_US_WR27_   INCTRPGNKI .RR.RIHIG. .PGRAFYTDR ..V..G...D IRQAYCNISG
B_US_YU2_M   INCTRPNNNT .RK.SINIG. .PGRALYTTG EII..G...D IRQAHCNLSK
BF1_BR_93B   INCTRPNNNT RKS..IPIG. .PGRAFYTTG EII..G...D IRKAHCNVSG
C_BR_92BR0   INCTRPNNN. .TRKSIRIG. .PGQAFYATG .....EIIGD IRQAHCNISR
C_BW_96BW0   IVCTRPGNN. .TRRSMRIG. .PGQTFYATG .....EIIGD IR.AHCNISE
C_BW_96BW1   IVCIRPNNN. .TRKSVRIG. .PGQTFYATE .....AIIGN ISEAHCNISE
C_BW_96BW1   IVCTRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGD IRQAHCNISK
C_BW_96BW1   IVCTRPNNN. .TRKGIRIG. .PGQTFYATE .....NIIGD IRQAHCNISA
C_ET_ETH22   ITCTRPSNN. .TRESIRIG. .PGQTFYATG .....DIIGD IRQAHCNISE
C_IN_93IN1   IVCTRPNNN. .TRKSIRIG. .PGQTFYATG .....DIIGD IRQAHCNISR
C_IN_93IN9   IVCTRPNNN. .TRKSIRIG. .PGQTFYATG .....EIIGD IRQAHCNISK
C_IN_93IN9   IECVRPNNN. .TRESIRIG. .PGQTFYATG .....EIIGD IRQAHCNISA
C_IN_94IN1   IVCTRPNNN. .TRKSIRIG. .PGQTFYATG .....EIVGN IRQAHCNISK
C_IN_95IN2   IMCTRPDNN. .TRKSIRIG. .PGQTFYATG .....DIIGD IRQAHCNISE
CRF01_AE_C   INCTRPFKN. .MRTSARIG. .PGQVFYKTG .....SITGD IRKAYCEING
CRF01_AE_C   INCTRPFKK. .VRISARIG. .PGRVFHTTG .....NINGD IRKAYCEINK
CRF01_AE_C   INCTRPFKK. .MRTSVRIG. .PGRVFYKTG .....SITGD IRKAYCEING
CRF01_AE_T   INCTRPSNN. .MRTSMRIG. .PGQVFYRTG .....SITGD IRKAYCEING
CRF01_AE_T   INCTRPSNN. .TRTSITMG. .PGQVFYRTG .....DIIGD IRKAYCEING
CRF01_AE_T   INCTRPFYN. .KRTRTSIG. .QGRVLYRTG .....DITGN IGKPYCEING
CRF01_AE_T   INCTRPSGK. .EEKRMTIG. .PGKVFYSTG .....KITGD IRKAYCVING
CRF01_AE_T   INCTRPSNN. .TRTSITIG. .PGRVFYRTG .....DIIGN IRKAYCEING
```

```
CRF01_AE_T  INCTRP.TI. .YKKKTTMG. .PARVYHRTG .....DVIGD IRKAYCQING
CRF02_AG_F  INCPRPNNN. .TRKSVRIG. .PGQTFYATG .....DIIGD IRQAHCNVSR
CRF02_AG_F  INCTRPNNN. .TRKSVRIG. .PGQTFYATG .....DIIGD IRKAHCNVSR
CRF02_AG_G  INCTRPNNN. .TRKSVRIG. .PGQTFYATG .....GIIGD IRQAHCNVSR
CRF02_AG_N  INCTRPNNN. .TRKGVHIG. .PGQAFYATG .....DIIGD IRQAHCNVSK
CRF02_AG_S  INCTRPGNN. .TRKSVRIG. .PGQTFYATG .....DIIGD IRQAHCNVSW
CRF02_AG_S  INCTRPSNN. .TRKSVRIG. .PGQTFYTTG .....AVIGD IRQAHCNVSR
CRF03_AB_R  INCTRPNNNT .RK.GIHIG. .PGRAFYATG DIT..G...D IRQAHCNISI
CRF03_AB_R  INCTRPNNNT .RK.GIHIG. .PGRAFYATG DII..G...D IRQAYCNISR
CRF04_cpx_  INCTRPGNNT RKS..VHIG. .PGLTWYATG EII..G...D IRQAHCNISG
CRF04_cpx_  INCTGLNNNT GGSERIGIG. .PGHTWYATG NIV..G...D IRQAHCNISG
CRF04_cpx_  INCTRPNNNT RKG..VHIG. .PGKTWFATG EVI..G...D IRKAHCNISE
CRF05_DF_B  INCTRPNNNT RKS..IHLG. .PGQAFYATG DII..G...D IRKAHCNVSR
CRF05_DF_B  INCTRPNNNT RKS..IPLG. .PGQAFYTTG DII..G...D IRKAHCNVSG
CRF06_cpx_  IRCTRPGNNT RKSISFGPG. ...QAFIATG DII..GD... IRQAHCNVSR
CRF06_cpx_  ISCSRPNNNT RKSIHIGPG. ...QAFYATG EII..GN... IRKAHCNVSR
CRF06_cpx_  IRCTRPSNNT RKSIPLGPG. ...QAFYATG DII..GD... IRQAHCNVSR
CRF06_cpx_  IKCTRPNNNT RKSISFAPG. ...QAFYATG DII..GD... IRQAHCNVSR
CRF11_cpx_  INCTRPGNNT RKSIHLGPG. ...HAFYATG AII..GD... IRQAHCKVSK
CRF11_cpx_  INCTRPNNNT RKGIHIGPG. ...QAFYATG DII..GD... IKQAHCNVSR
D_CD_84ZR0  INCTRPYKKE .RQ.RTPIG. .QGQALYTTR YT.....TRI IGQAYCNISG
D_CD_ELI_K  ITCARPYQNT .RQ.RTPIG. .LGQSLYTTR SR......SI IGQAHCNISR
D_CD_NDK_M  INCTRPYKYT .RQ.RTSIG. .LRQSLYTIT GKK..KKTGY IGQAHCKISR
D_UG_94UG1  INCIRPYNNT .RQ.STRIG. .PGQALFTTK VIG......D IRQAHCNISG
F1_BE_VI85  INCTRPNNNT RKG..IHLG. .PGQTFYATG AII..G...D IRKAHCNISG
F1_BR_93BR  INCTRPNNNT RKR..ISLG. .PGRVFYTTG EII..G...D IRKAHCNVSG
F1_FI_FIN9  INCTRPNNNT RKS..IRIG. .PGQSFYATG EII..G...D IRKAHCNISG
F1_FR_MP41  INCTRPNNNT RKS..IHLG. .PGQAFYATG DII..G...D IKKAYCEING
F2_CM_MP25  INCTRPNNNT RRS..IHIG. .PGRAFYATG EII..G...D TRKAHCNISE
F2KU_BE_VI  IVCIRPGNNT RKS..IRIG. .PGQTFYATG DII..G...D IRQAHCNITG
G_BE_DRCBL  INCTRPNNNT RRSVAIGPG. ...QAFYTTG EVI..GD... IRKAHCNVSW
G_NG_92NG0  INCIRPNNNT RKSIPIGPG. ...QAFYATG DII..GD... IRQAHCNVSR
G_SE_SE616  INCTRPNNNT MKRIRMGIGP ..GQTFYATG AII..GD... IRQAHCNVTK
H_BE_VI991  INCTRTGNNT RKS..IRIG. .PGQAFYATG DII..G...D IRRAYCNISG
H_BE_VI997  ITCTRPNNNT RKG..IHFG. .PGQAFYATG DII..G...N IRQAHCNVSE
H_CF_90CF0  ITCTRPNNNT RTS..IHLG. .PGRAFYATG DII..G...D IRQAHCNISR
J_SE_SE702  IVCTRPNNNT RKGIHMGPG. ...QVLYATG EII..GD... IRKAYCNISR
J_SE_SE788  IVCYRPNNNT RKGIHMGPG. ...QVLYATG EII..GN... IRETHCNISE
K_CD_EQTB1  INCTRPSNNT RKS..IHIG. .PGRAFYATG DII..G...D IRQAHCNISG
K_CM_MP535  INCTRPNNNT RKS..IHMG. .PGKAFYTTG DII..G...D IRQAHCNISG
N_CM_YBF30  INCTRPGNN. .TGGQVQIG. .PAMTFYNIE K.....IVGD IRQAYCNVSK
O_CM_ANT70  MTCERP.QI. .DIQEMRIG. ..PMAWYSMG IGG..TAGNS SRAAYCKYNA
O_CM_MVP51  MTCIREGIA. .EVQDIYTG. ..PMRWRSMT LKRSNNTSPR SRVAYCTYNK
O_SN_99SE_  MTCVRQGNQ. .SVQEIQIG. ..PMAWYSMS LAQE.GKPNN SRIAYCKYNI
O_SN_99SE_  MTCERPGNQ. .TVQKILTG. ..PVAWYSMG LKN...NLTN SRAASCKYNS
U_CD___83C  INCTRPGSDK KIRQSIRIG. .PGKVFYAKG ........GI TGQAHCNITD
```

```
            401                                                        450
00BW0762_1  TKWNKTLSRI GEKLKEHFPN ...KTIRFN. .SSAGGDLEI TTHSFNCRGE
00BW0768_2  EKWNSTLQGV REKLEKHFPN ...KNITFE. .PSSGGDLEI TTHSFNCRGE
00BW0874_2  QKWNKTLEQV GKKLAEHFPN ...KTIIFN. .SSSGGDLEI TTHSFNCGGE
00BW1471_2  SNWNSTLQQV ARKLEKYFPN ...KTIKFQ. .PSSGGDLEI TTHSFNCRGE
00BW1616_2  TKWENTLHMV SEKLKENFPN ...KTIVFN. .SSAGGDLEI TTHSFNCRGE
00BW1686_8  SKWNETLQKV KKKLGEHFPD ...KNITFE. .SPSGGDLEI TTHTFNCRGE
00BW1759_3  GEWNETLMEV SKELRKYFPN ...KNITFA. .PSSGGDLEI TTHSFNCRGE
00BW1773_2  AQWNKTLQEV GAKLEEHFPN ...KTIKFN. .QSSGGDLEI TTHSFICRGE
00BW1783_5  KAWNKTLHRV SEXLKEHFPN ...KTIKFT. .SPSGGDLEI IQHTFNCGGE
00BW1795_6  SKWITTLHRV SEKLKEHFPN ...KAINFT. .QPKGGDLEI TTHSFTCRGE
00BW1811_3  TNWNKTLQMV SEKLQQHFPN ...KTIKFD. .KHSGGDLEI TTHSFNCRGE
00BW1859_5  EEWNKTLQGV EEKLKEHFPN ...KKITFK. .PSSGGDLEV TTHSFNCGGE
00BW1880_2  DVWRKTLFNV SNKLKEYFPK ...RNITFN. .SSTGGDLEI TTHSFNCRGE
00BW1921_1  GAWNKTLQKV GEELRKHFPN ...KTIQFN. .SSSGGDLEI TTHTFNCRGE
00BW2036_1  SAWNRTLHRV SKKLREHFPN ...TTIKFQ. .PASGGDLEI TTHSFNCRGE
00BW2063_6  EKWNKTLYRV SEKLKEYFPN ...KTIKFD. .QPTGGDLEI TKHSFNCRGE
00BW2087_2  DKWNKTLQQV GEKLAEHFPN ...KTIKFA. .PSSGGDLGI TTHSFNCRGE
00BW2127_2  DAWNETLQQV GKKLEEYFPN ...KTIKFA. .NSSGGDLEI TTHSFNCGGE
00BW2128_3  EEWNKTLREV KGKLGKHFN. ...KTIMFA. .PSSGGDLEI TAHSFNCRGE
00BW2276_7  NQWNETLQRV GKKLKEHFN. ...KTIKFE. .QSSGGDLEI TMHSFNCRGE
00BW3819_3  KAWNKTLQLV VKKLKEHFPN ...RTIKFT. .PPPGGNLEI TTHSFNCRGE
00BW3842_8  GNWTKTLQRV SRKLRQIFNK ...SNIEFK. .PHSGGDPEI TTHSFNCRGE
00BW3871_3  HEWNKTLEGV KKKLGEHFPN ...KTIKFQ. .PSSGGDLEI ATHTFNCRGE
00BW3876_9  SDWNRTLQWV KNKLGEHFPN ...TVIKFE. .PSSGGDLEI TTHSFNCRGE
00BW3886_8  GNWTETLRRV KNKLGEHFPN ...KTITFG. .PSSGGDLEI TTHSFNCRGE
00BW3891_6  STWNRTLQEV SRKLVERFPN ...KTIRFQ. .PPSGGDLEI TTHSFNCRGE
00BW3970_2  ADWNRTLQGV GRKLAGYFPN ...KTISFQ. .PSSGGDLEI TTHSFNCGGE
00BW5031_1  TKWTETLQKI SEKLRGYFK. ...KTIIFA. .PSSGGDPEI TTHSFNCRGE
 96BW01B21  EVWKKTLQRV GRKLKEHFPN ...KTIQFQ. .PPSGGDLEV TTHSFNCRGE
  96BW0407  KDWNKTLHRV RKKLAEHFS. ...KNITFK. .PSSGGDLEI TTHSFNCGGE
  96BW0502  TEWNSTLQGV SKKLEEHFSK ...KAIKCE. .PSSGGDLEI TTHSFNCRGE
 96BW06_J4  TNWNKTLKGV KGELRKHFH. ...KNITFQ. .PASGGDLEI TTHTFNCRGE
 96BW11_06  SQWNKTLHRV IEKLKEHFPN ...KTIGFS. .QAAGGDLEI TTHSFNCGGE
  96BW1210  GAWNETLQWV GKKLKEHFPN ...KTIRFK. .ESSGGDLEI TTHSFNCGGE
 96BW15B03  GEWNKAVQRV SAKLREHFPN ...KTIEFQ. .PSSGGDLEI TTHSFNCRGE
 96BW16_26  SEWKRTLQRV SEKLGKHFPN ...KTIKFA. .PHSGGDLEI TTHVLNCRGE
 96BW17A09  ..WNNTLQQV VKKLRKHFPN ...KTITFA. .PPSGGDLEI TTHSFNCGGE
 96BWMO1_5  SKWNKILYRV SEKLKEHFPN ...KTIQFG. .QPIGGDLEI TTHSFNCRGE
 96BWMO3_2  GEWAKVMQKV TGKLKEHFP. ...KNITFQ. .PPSGGDLEI TTHSFNCRGE
98BWMC12_2  QKWNKTLQQV GKKLAEYFPN ...RTIAFN. .SSARGDLEI ATHSFNCRGE
98BWMC13_4  SLWNETLYKV SEKLKEYFN. ...TTIEFQ. .QPAGGDLEI TTHSFNCRGE
98BWMC14_a  EEWNNSLQKV AKKLREHFPN ...KTIAFN. .SSSGGDLEI TTHTFNCRGE
98BWMO14_1  GNWTKTLHRV GEKLKEHFPN ...KTIKFA. .PPSGGDLEI IMHSFNCGGE
98BWMO18_d  DDWKSTLQNV SEKLRRHFPN ...KTIVFN. .SPSGGDLEI ITHSFNCEGE
98BWMO36_a  ERWNKTLQEV GEKLERHFPN ...KTIEFK. .PSSGGDLEI TTHSFNCRGE
98BWMO37_d  KKWYKALHGV REKLKVLFPN ...KNISFQ. .PAAGGDLEV TTHSFNCKGE
99BW3932_1  DAWNKTLQQV GRKLKEYFPG ...STIRFA. .PHSGGDLEI TTHSFNCRGE
99BW4642_4  SEWNRTLQRV GEKLGKYFPS ...KTIKFN. .SSSGGDLEI TTHSFNCRGE
99BW4745_8  KAWNTTLQEV GRKLAEHFPN ...KTIKFQ. .PHSGGDPEI TMHSFTCGGE
99BW4754_7  SQWNKTLYEV SEKLREKFPN ...KTIQFN. .SSTGGDLEI TTHSFNCGGE
99BWMC16_8  EAWNKTLLRV SKKLREHFPN ...KTIIFD. .KSSGGDLEI TTHSFNCGGE
A2_CD_97CD  TEWNATLKKV VEQLREHFPN ...KTIIFN. .SSSGGDLEI TTHSFNCGGE
A2_CY_94CY  TLWNDTLQKV AEQLREKFPK ...KTIIFT. .NSSGGDPEI TTLSFNCAGE
A2D___97KR  RQWNDTLQKV AEQLRKYFSN ...KTIIFT. .NSSGGDIEI TTHSVNCGGE
A2G_CD_97C  AGWNDILQKV AEQLGKHFSG ...KNITFA. .NSSGGDLEI TTHSFNCGGE
A_BY_97BL0  AAXNSTLQKI STQLRKYFNN ...KTIIFK. .SSTGXDLEV TTHSFNCGGE
A_KE_Q23_A  SRWNKTLQEV AEKLRTYFGN ...KTIIFA. .NSSGGDLEI TTHSFNCGGE
A_SE_SE659  SEWNNTLQQV AKQLRTYFGN ...KTIIFT. .NSSGGDLEI TTHSFNCKGE
A_SE_SE725  SSWNKTLQDI VTQLRVYWN. ...RTIIFN. .SSSGGDLEI TTHSFNCGGE
A_SE_SE753  SKWNATLQKV AIKLREYFDD ...KTIIFT. .KPSGGDLEI TTHSFNCGGE
```

```
A_SE_SE853   AKWNKTLHEV AKQLRTYFNN ...KTIIFT. .NSSGGDLEI TTHTVNCGGE
A_SE_SE889   TEWNEALQKV VNQLKTHFKN ...KTIIFN. .SSSGGDLEI TTHSFNCGGE
A_SE_UGSE8   SKWNETLKKV AIQLRKYWN. ...TTIIFT. .NSSGGDLEI TTHSFNCGGE
A_UG_92UG0   SQWNKTLHQV VEQLRKYWNN ...NTIIFN. .SSSGGDLEI TTHSFNCAGE
A_UG_U455_   RDWNRTIQQV AEQLKKKFNN ...KTIIFA. .SSSGGDIEI TTHSFNCGGE
AC_IN_2130   AEWNKALNKI GKQLRKYFVN ...KTIKFA. .NSSGGDLEI TTHSFNCEGE
AC_RW_92RW   TKWNRTLQKV AEKLSHYFEN ..ITTIIFK. .NSSGGDLEI TTHSFNCGGE
AC_SE_SE94   TKWNKTLHKV VTQLRKYFVN ...KPIIFT. .PSSGGDVEV TTHSFNCRGE
ACD_SE_SE8   SEWNKTLQQV AKKLGDPLNK ...TEIIFK. .PPSGGDLEI TTHSFNCGGE
ACG_BE_VI1   KEWNKTLQAV GKKLAEYYPN ...KTINFT. .QASGGDLEI VTHSFNCGGE
AD_SE_SE69   TKWDKTLREV AIQLKHYYG. ..NKTVIFAN .SS.GGDIEI TTHSFNCRGE
AD_SE_SE71   SAWNNTLQQV VIQLRRYFNN ...KTIIFT. .NSSGGDLEI TTHSFNCGGE
ADHK_NO_97   GSWMKTLHKV ATQLXQHFS. ..NKTIIFNA .SA.GGDIEI TTHSFNCAGE
ADK_CD_MAL   TEWDKTLQQV AVKLGSLLN. ..KTKIIFNS .SS.GGDPEI TTHSFNCRGE
AG_BE_VI11   KDWGKMLQEV SRQLKKFFNN ...KTIFFNS .SA.GGDLEI TTHSFNCRGE
AG_NG_92NG   QEWQEMLQKV QAQLEQVFN. ...KSITFNS .SA.GGDLEI TTHSFNCRGE
AGHU_GA_VI   EQWNRTLERV KEKLGRHFK. ..NKTITFKP .AS.GGDPEV TMHIFNCRGE
AGU_CD_Z32   KEWSETLSKV AAQLRKHFVN T.RTDIIFA. .NSSGGDVEI TTHSFNCGGE
AJ_BW_BW21   KNWTDTLHKV TAKLKEYFN. ...TTIEFQP .AS.AGDLEI MTHTFNCGGE
B_AU_VH_AF   TNWTSVLRQI AVKLRERFK. ..NKTIVFNH .SS.GGDPEI VRHSFNCGGE
B_CN_RL42_   TKWNNTLKQI TKKLREQFG. ..NKTIVFNQ .SS.GGDPEI VMHSFNCGGE
B_DE_D31_U   AKWDSTLRQI VKKLRERFG. ..NKTIVFNQ .SS.GGDPEI VTHSFNCGGE
B_DE_HAN_U   ARWNKTLNQI FRKLREIRQF .ENKTIVFNR .SS.GGDPEI VMHSFNCGGE
B_FR_HXB2_   AKWNNTLKQI ASKLREQFG. .NNKTIIFKQ .SS.GGDPEI VTHSFNCGGE
B_GA_OYI__   ATWEKTLEQI ATKLRKQFR. .N.KTIAFDR .SS.GGDPEI VMHSFNCGGE
B_GB_CAM1_   TKWNNTLKQI VTKLKEQFG. ..NKTIIFNQ .SS.GGDPEI VMHSFNCGGE
B_GB_GB8_C   EKWNNTLHQI VIELRKQFR. ..NKTIVFNQ .SS.GGDPEI VMHSFNCGGE
B_GB_MANC_   AKWEKTLKQI VEKLREKFG. ..NKTIIFNQ .SS.GGDPEI VTHSFNCGGE
B_KR_WK_AF   EKWNDTLKQL VIKLGEQFG. .NSNIIVFKQ .SS.GGDPEI VMHSFICGGE
B_NL_3202A   AKWNNTLKQI VSKLRKQFG. ..NKTIVFSQ .PL.GGDPEI VMHSFNCGGE
B_TW_TWCYS   AEWNNTLPQI VKKFREQFG. ..NKTIVFNQ .SS.GGDLEI VMHSFNCGGE
B_US_BC_L0   AKWNDTLRQI VIKLR..EQF .ENKTIVFNQ .SS.GGDPEI VMHSFNCGGE
B_US_DH123   VKWHNTLKRV VEKLREKFE. ..NKTIVFNK .SS.GGDPEI VMHSFNCGGE
B_US_JRCSF   AQWNNTLKQI VEKLR..EQF .NNKTIVFTH .SS.GGDPEI VMHSFNCGGE
B_US_MNCG_   AKWNDTLRQI VSKLKEQFK. ..NKTIVFNQ .SS.GGDPEI VMHSFNCGGE
B_US_P896_   AKWNNTLQQI VIKLR..EKF .RNKTIAFNQ .SS.GGDPEI VMHSFNCGGE
B_US_RF_M1   AQWNNTLKQV VTKLR..EQF .DNKTIVFTS .SS.GGDPEI VLHSFNCGGE
B_US_SF2_K   AQWNNTLEQI VKKLREQFG. .NNKTIVFNQ .SS.GGDPEI VMHSFNCRGE
B_US_WEAU1   TSWNNTLKQI VEKLREIKQF .KNKTIVFKQ .SS.GGDPEI VMHSFNCGGE
B_US_WR27_   TKWKNTLEKI VAKIREIKQF .KNKTIVFNH .SS.GGDPEI VMHSFNCGGE
B_US_YU2_M   TQWENTLEQI AIKLKEQFG. .NNKTIIFNP .SS.GGDPEI VTHSFNCGGE
BF1_BR_93B   TKWNETLEKV RAKLKPHFPN ...ATIKFNS .SS.GGDLEI TMHSFNCRGE
C_BR_92BR0   TAWNKTLQEV GKKLAEHFPN ...KAIKFA. .KHSGGDLEI TTHSFNCRGE
C_BW_96BW0   RDWNDTLNRV SKKLAEHFPN ...KTIEFK. .PSSGGDLEI TTHSFNCRGE
C_BW_96BW1   SQWNNTLQRV SEKLKEHFPN ...KTIKFN. .QPAGGDLEI TTHSFNCGGE
C_BW_96BW1   GAWNETLQWV GKKLKEHFPN ...KTIRFK. .ESSGGDLEI TTHSFNCGGE
C_BW_96BW1   GEWNKAVQRV SAKLREHFPN ...KTIEFQ. .PSSGGDLEI TTHSFNCRGE
C_ET_ETH22   EKWNKTLQKV KEKLQKHFPN ...KTIEFK. .PSSGGDLEI TTHSFNCGGE
C_IN_93IN1   DKWNETLQRV GKKLAEHFHN ...KTIKFA. .SSSGGDLEI TTHSFNCRGE
C_IN_93IN9   ENWTDTLQRV SKKLAEHFPN ...KTIKFD. .SPSGGDLEI TTHSFNCRGE
C_IN_93IN9   DRWNETLQWV GEKLAEHFPN ...KTIKFA. .PSSGGDLEI TTHSFNCRGE
C_IN_94IN1   RDWNETLQRV SEKLAKHFPN ...KTIKFA. .PSSGGDLEI TTHSFNCRGE
C_IN_95IN2   DKWNETLQNV SKKLAEHFPN ...KTIIFN. .SSSGGDLEI TTHSFNCRGE
CRF01_AE_C   TKWNETLKQV TKKLREHFKN ...KTIIFQ. .PSSGGDPEI TMHHFNCRGE
CRF01_AE_C   TKWKETLKQV TRKLREHLNG ...TMTISFR .PSSGGDPEI TMHHFNCRGE
CRF01_AE_C   TKWNETLQQI IRKLEEHFNN ...KTIQFKP .PYSGGDLEI TMHHFNCRGE
CRF01_AE_T   TKWNKVLKQV TEKLKEHFNN ...KTIIFQ. .PPSGGDLEI TMHHFNCRGE
CRF01_AE_T   TKWNEVLKQV AGKLKEHFNN ...KTIIFK. .PPSGGDLEI TMHHFNCRGE
CRF01_AE_T   TKWNKVLNQV TEKLKEHFNN ...RNISFQ. .PPSGGDLEI TMHHFICRGE
CRF01_AE_T   TKWNETLKQV AGKLREHFNN ...KTIIFQ. .PPSGGDLEI TMHHFNCRGE
CRF01_AE_T   TKWNKVLKQV TEKLKEHFN. ...KTIIFQ. .PPSGGDLEI TMHHFNCRGE
```

```
CRF01_AE_T  TKWNKVLKQV TEKLKEHFN. ...KTIIFQ. .PPSGGDLEI TMHHFNCRGG
CRF02_AG_F  SEWNRTLQQV ATQLRKHFN. ...KTIIFA. .NSSGGDIEI TTHSFNCGGE
CRF02_AG_F  SKWNNTLQQV AIQLRKHFN. ...TTIIFA. .NPSGGDIEI TTHSFNCGGE
CRF02_AG_G  TDWNTTLQQV ATQLGKYFRD T..TRIKFD. .NPSGGDLEI MTHSFNCGGE
CRF02_AG_N  TEWNKTLHQV VTQLKTYFKN ...TTIIFA. .NPLGGDVEI TTHSFNCGGE
CRF02_AG_S  QQWNKTLHDV ATKLREYFNN ...TTIIFD. .EPSGGDLEI TTHSFNCGGE
CRF02_AG_S  EKWNSTLQKV VTKLGKHFNS ...SKIIFT. .NSSGGDLEI TTHSFNCGGE
CRF03_AB_R  TKWNNTLKQI VIKLRKQFG. ..NKTIVFNQ .SS.GGDPEI VMHSFNCGGE
CRF03_AB_R  TKWNNTLEQI VSKLRKQFR. ..NKTIVFNQ .SS.GGDPEI VMHSFNCGGE
CRF04_cpx_  NDWNDTLKVI SEELKRLFP. ..NKTIKFAP .PV.GGDLEI TTHSFNCKGE
CRF04_cpx_  SDWNEALQKV VVKLREHFP. ..NKTIIFNQ .SS.GGDLEI TTHSFNCGGE
CRF04_cpx_  KDWNTTLQKI VDELRKHFP. ..NKNITFAP .SA.GGDVEI TTHSFRLGGE
CRF05_DF_B  EQWNKTLIQV AKELQSHFP. ..NKTIKFNS .SS.GGDLEI TMHSFNCRGE
CRF05_DF_B  AQWNKTLEQV KEELRAHIKD IGNKTIVFNS .SA.GGDLEI TSHIFNCRGE
CRF06_cpx_  ANWTDILGEV KVKLEEVFNN ...THITFKS .SA.GGDLEI TTHSFNCGGE
CRF06_cpx_  KAWNSMLQNV TAKLKELFNN ...KNITFNS .SA.GGDLEV TTHSFNCGGE
CRF06_cpx_  TAWKETLQNV TEKLKQLLN. ...TNITFNP .SA.GGDLEI TTHSFNCRGE
CRF06_cpx_  TDWNNMLKNV TTKLIEVFK. ...KNITFNS .SA.GGDLEI TTHSFNCGGE
CRF11_cpx_  AEWLNTLQQV ATQLRGKFN. ...KTIIFDN .PSPGGDIEI TSHSFNCRGE
CRF11_cpx_  ADWNNTLQQV AEQLHNNFN. ...KTIVFNE .HS.GGDLEV TTHSFNCGGE
D_CD_84ZR0  VKWNNTLRQV ARKLGNLLN. ..QTKIIFKP .SS.GGDPEI TTHSFNCGGE
D_CD_ELI_K  AQWSKTLQQV ARKLGTLLN. ..KTIIKFKP .SS.GGDPEI TTHSFNCGGE
D_CD_NDK_M  AEWNKALQQV ATKLGNLLN. ..KTTITFKP .SS.GGDPEI TSHMLNCGGD
D_UG_94UG1  AGWNKTLQQV AEKLGNLLN. ..QTTIIFKP .SS.GGDPEI TTHSFNCGGE
F1_BE_VI85  TQWNNTLEYV KAELKSHFPN N..TAIKFNQ .SS.GGDLEI TMHSFNCRGE
F1_BR_93BR  TQWRNTLAKV KAKLGSYFPN ...ATIKFNS .SS.GGDLEI TRHNFNCMGE
F1_FI_FIN9  EQWNKTLDRV KAELKLHFNK ....TIQFNS .SS.GGDLEI TMHSFNCRGE
F1_FR_MP41  TQWSKTKTQV QEKLRALFNK ....TIKFNQ .SS.GGDLEI TMHSFNCRGE
F2_CM_MP25  KQWYDTLIKI ATEFKDQYN. ...KTVGFQP .SA.GGDLEI TTHSFNCRGE
F2KU_BE_VI  ENWNKTLEGV KAKLHGFFTN ...KTIIFKP .HS.GGDPEV VMHTFNCGGE
G_BE_DRCBL  TKWNETLRDV QAKLQEYFIN ...KSIEFNS .SS.GGDLEI TTHSFNCGGE
G_NG_92NG0  IKWREMLKNV TAQLRKIYN. ..NKNITFNS .SA.GGDLEI TTHSFNCRGE
G_SE_SE616  RKWKEALQNV AAELGKIFNK S.SENITFNS .SA.GGDLEI TTHSFICRGE
H_BE_VI991  KQWNETLHKV ITKLGSYFD. ..NKTIILQP .PA.GGDIEI ITHSFNCGGE
H_BE_VI997  EKWNKTLQQI ATQLSKYFV. ..NRTLIFKP .HS.GGDLEV TTHSFNCRGE
H_CF_90CF0  TDWNKTLHQV VTQLGIHLN. ..NRTISFKP .NS.GGDMEV RTHSFNCRGE
J_SE_SE702  KDWNNTLRRV AKKLREHFN. ...KTIDFTS .PS.GGDIEI TTHSFNCGGE
J_SE_SE788  RDWSNTLRRV ATKLREHFN. ...KTINFTS .PS.GGDIEI VTHSFNCGGE
K_CD_EQTB1  GQWNKTVNQV KKELGKHFN. ...KTIIFQP .SS.GGDPQV TRHIFNCRGE
K_CM_MP535  EKWNMTLSRV KEKLKEHFKN ...GTITFKP .PNPGGDPEI LTHMFNCAGE
N_CM_YBF30  ELWEPMWNRT REEIKKILGK ...NNITFRA RERNEGDLEV THLMFNCRGE
O_CM_ANT70  TDWGKILKQT AERYLELVNN TGSINMTFN. .HSSGGDLEV THLHFNCHGE
O_CM_MVP51  TVWENALQQT AIRYLNLVNQ TENVTIIFS. .RTSGGDAEV SHLHFNCHGE
O_SN_99SE_  SDWEKALKQT AERYLDLRNN TNTVNITFE. .RSIGGDSEV THLHFNCHGE
O_SN_99SE_  SVWEEALKQT AERYLELMNN TNTVNITFN. .HSTGGDPEV THLHFNCHGE
U_CD___83C  GEWRNTLQQV AIALRRQFNN ...KSIIFN. .SSSGGDIEI TTHTFNCGGE
```

```
           451                                                    500
00BW0762_1 FFYCNTTRLF NGTYN..... ......STGD TNS.....TN STITLQCRIK
00BW0768_2 FFYCDTSNLF NKTRR..... ......DN.. ........AN ETITLPCRIK
00BW0874_2 FFYCNTSRLF NSTYN..... ......PNST YIEGR...SN ATITLQCRIK
00BW1471_2 FFYCYTTKLF NSTYN..... ......STYT GSESN..... ..ITIPCRIK
00BW1616_2 FFYCNTSKLF NGTYN..... ......SNNN TA........ .DITLQCRIK
00BW1686_8 FFYCNTSNLS NETYL..... ......ANLT SNVTK....N ATITLPCRIK
00BW1759_3 FFYCNTSNLF NNTYR..... ......ADNN ITNDNSN... ..ITLQCRIK
00BW1773_2 FFYCNTSALF NSTYN..... ......STNT SGHN....DT RIITLPCRIK
00BW1783_5 FFYCNTSKLF NGTYN..... ......GTS. ...ISS...N SSITLQCRIK
00BW1795_6 FFYCNTSELF NGTYN..... ......STG. ...DSN...S NLITLQCRIK
00BW1811_3 FFYCNTSQLF NGTYM..... ......PNTY MS....SSDN RNITIPCRIK
00BW1859_5 FFYCNTTHLF NGNG...... .......... .......ESD INITLPCRIK
00BW1880_2 FFYCDTTKLF NGTYN..... ......STEQ TN........ STITLQCRIK
00BW1921_1 FFYCNTSQLF NGTYN..... ......DT.Y ESNSG....N STITLPCRIK
00BW2036_1 FFYCDTSKLF NSSYN..... ......DTEL YSYNS....T ANITLPCRLK
00BW2063_6 FFYCNTSQLF NSSYS..... ......RHN. ...NTS...N STITLPCNIK
00BW2087_2 FFYCNTSGLF N......... .......GTF NGT...HSTN TNITLPCRIK
00BW2127_2 FFYCNTTILF NSTYY..... ......P... NTK...SDTT ETITLPCRIK
00BW2128_3 FFYCNTSLLF DETQL..... ......SKE. .........N NTINIQCRIK
00BW2276_7 FFYCNTSKLF NGTYM..... ......PNYN TSN...SSNN SNITLPCRIK
00BW3819_3 FFYCNTSGLF NGTYN..... ......G... TND...NDTD SDITLPCKIK
00BW3842_8 FFYCNTSLLF NSSYN..... ......GNSS YNDTGS...N STITLQCRIK
00BW3871_3 FFYCNTSILF NDTYW..... ......FNGT ANDTG....S NNITIPCRIK
00BW3876_9 FFYCNTSGLF NNNLI..... ......NNG. ........AE DTIRLPCRIK
00BW3886_8 FFYCNTSKLF NSTNN..... ......NTE. ..SES....N ATITLPCRIK
00BW3891_6 FFYCNISRLF NRPNM..... ......TKNM TSDIKNN... STITLPCRIK
00BW3970_2 FFYCNTSSLF NNTYR..... ......PTYW PGTE....SN STITLQCRIK
00BW5031_1 FFYCNTSQLF NSTYR..... ......ANTS NS........ .NITLPCRIK
 96BW01B21 FFYCDTSELF NSTYM..... ......SNGG NISS.....S TIIMLPCRIE
  96BW0407 FFYCNTSRLF NESYN..... ......FDES YWN.N...TN KTIMLPCRIK
  96BW0502 FFYCDTSQLF NSTYS..... ......PSNG TENK....LN GTITITCRIK
 96BW06_J4 FFYCNTSRLF DETYL..... ......S... GTDED....N GTITLPCKIK
 96BW11_06 FFYCNTSKLF NSTYI..... ......QLN. .STETP...N STITLPCRIK
  96BW1210 FFYCNTSQLF NSTYN..... .........Y MPS...NNTG TNITLQCRIK
 96BW15B03 FFYCNSSKLL NSSYN..... ......GTSY RGTESN...S SIITLPCRIK
 96BW16_26 FFYCNTSKLF NSTYN..... ......STDR SNN......T DNITIQCRIK
 96BW17A09 FFYCNTSILF NSTYN..... ......STYT GSDSNS.... .TITIPCRIK
 96BWMO1_5 FIYCNTSKLF NGTYN..... ......STG. ....TS...N STITLSCRIK
 96BWMO3_2 FFYCNTSELF NGTYN..... ......GTD. ..NNS....N KTITLLCRIK
98BWMC12_2 FFYCNTSGLF NSTYN..... ......PNST YTESK...AN SNITLHCRIK
98BWMC13_4 FFYCNTTKLF NGTYS..... ......QPN. .STGTP...H SNITLPCKIK
98BWMC14_a FFYCNTSQLF NSTYN..... ......G... RNSTT....N ATITLPCRIK
98BWMO14_1 FFYCNTSKLF NSTYN..... .......ATY NST...DTSN STITIPCRIK
98BWMO18_d FFYCNTSGLF NS........ .......... AFNDN...SG GTITLQCRIE
98BWMO36_a FFYCNTSGLF NSTYY..... ......SNKT SSN...MTTN EIITIPCKIK
98BWMO37_d FFYCNTSKLF NTSWL..... ......DSYI SNTG....NN SIITLPCRIK
99BW3932_1 FFYCNTSRLF NSTYN..... ......P... NTK...SNTG SWIILPCRIK
99BW4642_4 FFYCNTSKLF TYQSN..... ......TY.. .......VAN STITLPCKIK
99BW4745_8 FFYCNTSELF NSTYN..... ......ANTY NTATGNNS.. TTIILPCRIK
99BW4754_7 FFYCNTSKLF NSTFN..... ......SNGH DST.....GN DPLTIPCRIK
99BWMC16_8 FFYCNTSNLF NNTYY..... ......PNMT NTDTK...SN LTITLPCRIK
A2_CD_97CD FFYCNTTGLF NSTWEN.... ......GTNK QNYTE...SN DTITLQCRIK
A2_CY_94CY FFYCNTTGLF NGTWWNN... ......GTWN GPYTPNN.TN GSIILPCRIK
A2D___97KR FFYCDTSGLF NSTWPAN... ......ASRE NEEKD...R. .NVTLPCRIK
A2G_CD_97C FFYCNTTNLF NSTFNTT... ......SLFN STGRNGTNDN TTITIPCRIK
A_BY_97BL0 FFYCNTTDLF NSTX...... .....DGTVT NSTKAN.... GTITLPCRIK
A_KE_Q23_A FFYCNTSGLF NSTWY..... .....VNSTW NDTDSTQESN DTITLPCRIK
A_SE_SE659 FFYCNTSSLF NSTWS..... .....NDNNT QGSNSTET.K GTITLPCRIK
A_SE_SE725 FFYCNTSGLF NSTWS..... .....Q.NDT GVSNSTES.N DTIILPCRIK
A_SE_SE753 FFYCNTSGLF NSTIL..... .....NSTKM NDNASRESYD DTITLQCRIK
```

```
A_SE_SE853   FFYCNTSGLF NSTWS..... .....SNASE PMSNSTES.N DTITLQCRIR
A_SE_SE889   FFYCNTSGLF NSTWN..... .....GTDSM QKLNST.... GNITLPCRIK
A_SE_UGSE8   FFYCNTSGLF NSSWN..... .....END.T KVNYNTES.N DTITLQCRIK
A_UG_92UG0   FFYCNTSGLF NSTWV..... .....NGTTS STSN...... GTITLPCRIK
A_UG_U455_   FFYCNTSGLF NSIWN..... .....GSMSN DMGP.....N GTITLQCRIK
AC_IN_2130   FFYCNTSGLF NGTWNASMQ. .........ES NSTESN.... ETIILPCRIK
AC_RW_92RW   FFYCNTSGLF NSTWS..... .........KR NGTWQSNGTE LNITLPCRIK
AC_SE_SE94   FFYCDTSGLF NSTWPFNS.. ..........T NSTGPN.... GTITLQCRIK
ACD_SE_SE8   FFYCNTSGLF NSTWV..... .....NGSRE SNSTDN.... DTITLPCRIK
ACG_BE_VI1   FFYCNTSGLF NSTYN..... ......PSYN STESVN...E TTIILPCKIK
AD_SE_SE69   FFYCNTTGLF NSTWNDTAT. ......EQKP .......N.. DTIRLQCRIK
AD_SE_SE71   FFYCNTSGLF NSTWN..... .....NTDSM QESHSTET.N DTITLPCRIK
ADHK_NO_97   FFYCNTSQLF NSTWNHTST. ......YNST EN........ GTITLPCKIK
ADK_CD_MAL   FFYCNTSKLF NSTWQNNGA. ......RLSN S..TE.ST.. GSITLPCRIK
AG_BE_VI11   FFYCNTSALF NFSSETNST. .......... FP.N...... TTLTLPCRIK
AG_NG_92NG   FFYCNTSGLF NESGGNDT.. .......... .......... .TITLPCKIK
AGHU_GA_VI   FFYCNTTKLF NDTENKN... .......... ..NDAEN... KTITLPCRIR
AGU_CD_Z32   FFYCNTSGLF NSTWK..... .....NSTSI NDTVSN.... GTITLPCRIK
AJ_BW_BW21   FFYCNTSGLF NKSLLNETS. ......NETT DGAN...... NTITLTCRIK
B_AU_VH_AF   FFYCNSTQLF NSTWFNSTG. ......NDTE RATNN..T.. ENITLPCRIK
B_CN_RL42_   FFYCNTSQLF NSTWNDTG.. .........T WNDTTGNS.. .TITLPCRIK
B_DE_D31_U   FFYCNSAQLF NSTWNDTK.. .........ES NNTNG..... .TITLPCRIK
B_DE_HAN_U   FFYCNSTKLF NSTWNNTST. ......WN.. DNGND..... .TIILPCRIK
B_FR_HXB2_   FFYCNSTQLF NSTWFNSTW. ......STEG SNNTEGSD.. .TITLPCRIK
B_GA_OYI__   FFYCNTSQLF NSTWNDTTR. ......AN.. .STEV..... .TITLPCRIK
B_GB_CAM1_   FFYCNTTQLF NTTWLFNGT. ......WNDT EGLNNTER.. .NITLPCRIK
B_GB_GB8_C   FFYCKTAQLF NSTWNSTGN. ......GTIK SNTTE..... .IITLPCRIK
B_GB_MANC_   FFYCNSTQLF NSTWNTGND. ......TRES NDTNN..T.. GNITLPCRIK
B_KR_WK_AF   FFYCNTTQLS NSTWQRSDG. ......TWNR TGGLNETK.. ENITLPCRIK
B_NL_3202A   FFYCNSTQLF NSTWNDTGN. ......VTER SNNNE..... .NITLPCRIK
B_TW_TWCYS   FFYCNATPLF NSTWNATST. ......LNAT NEENE..... .NITLLCRIK
B_US_BC_L0   FFYCKSTQLF NSTWAGNNT. ......WNSS AERSDDTG.. GNITLPCRIK
B_US_DH123   FFYCNTKKLF NSTWNGTEG. ......SYNI EGND...... .TITLPCRIK
B_US_JRCSF   FFYCNSTQLF NSTWNDTEK. ......SSG. TEGND..... .TIILPCRIK
B_US_MNCG_   FFYCNTSPLF NSTWNGNNT. ......WNNT TGSNN..... .NITLQCKIK
B_US_P896_   FFYCNTAQLF NSTWNVTGG. ......TNG. TEGND..... .IITLQCRIK
B_US_RF_M1   FFYCNTTQLF NSTWNSTEG. ......SNNT GGND...... .TITLPCRIK
B_US_SF2_K   FFYCNTTQLF NNTWRLNHT. ......EG.. TKGND..... .TIILPCRIK
B_US_WEAU1   FFYCNSTQLF NSTWHANGT. ......WKNT EGADN..... .NITLPCRIK
B_US_WR27_   FFYCNSTQLF NSTWNSTEG. ......NS.. TWSDK..... .IIRLPCRIK
B_US_YU2_M   FFYCNSTQLF ..TWNDTRK. ......LN.. .NTGR..... .NITLPCRIK
BF1_BR_93B   FFYCNTSGLF NDTVDN.... .......... .......... GTITLPCRIK
C_BR_92BR0   FFYCNTSSLF NSTYT..... ......PNST ENITGT..EN SIITIPCRIK
C_BW_96BW0   FFYCNTSRLF NESYS..... ......FNES HWSND...TN ATITLPCRIK
C_BW_96BW1   FFYCNTSKLF NGTYI..... ......QPNS .TEDTP...N STITLPCRIK
C_BW_96BW1   FFYCNTSQLF NSTYN..... ......S.TY MPS...NNTG TNITLQCRIK
C_BW_96BW1   FFYCNSSKLL NSSYN..... ......GTSY RGTESN...S SIITLPCRIK
C_ET_ETH22   FFYCNTSNLF NSTKL..... ......E... LFNSS...TN LNITLQCRIK
C_IN_93IN1   FFYCNTSGLF NGTYM..... ......PTYM PNGTESN.SN STITIPCRIK
C_IN_93IN9   FFYCNTSGLF NGTYN..... ......TSSD GNS......S STITIPCRIK
C_IN_93IN9   FFYCNTSSLF DSLFN..... ......PNGT RNDT.....N LTITIPCRIK
C_IN_94IN1   FFYCNTSGLF NSTYM..... ......SGTY MNSSADM.NS SYITIPCRIK
C_IN_95IN2   FFYCNTSGLF NRTYM..... ......PNDT KSNSSSN.PN ANITIPCRIK
CRF01_AE_C   FFYCNTTKLF NSTWT..... .......TNE IMEEFKGTNS STITLPCRIK
CRF01_AE_C   FFYCNTTALF NSTWI..... .......N.G TMQEVNGTNS GNITLPCRIK
CRF01_AE_C   FFYCNTTRLF N......... .......... ..ISTNGTTN GTITLPCRIK
CRF01_AE_T   FFYCNTTQLF NNTCI..... .......GNE TMK...GCNG .TITLPCKIK
CRF01_AE_T   FFYCNTTQLF NSTWT..... .......GNE TME...GSNG .TITLPCKIK
CRF01_AE_T   FFYCNTTRLF NNTCI..... .......GNK TMK...ECND .TIILPCKIK
CRF01_AE_T   FFYCNTTKLF NSTWI..... .......GNE TIG....SSG .NIILPCRIK
CRF01_AE_T   FFYCNTTKLF NNTCL..... .......GNE TMA...GCND .TITLPCKIK
```

```
CRF01_AE_T  FFYCNTTKLF NSTWR..... .......GNE TIESREGYNK .TIILPCKIK
CRF02_AG_F  FFYCNTSELF NSTW...... ..NSTWDNSS NHIESNHT.E GNITLQCRIK
CRF02_AG_F  FFYCNTSELF N......... ...STWDNSL NHTESNHT.E DNITLQCRIK
CRF02_AG_G  FFYCNTSGLF NSTWYKN... ..STWYSNST ASSNHTEL.N STITLQCKIK
CRF02_AG_N  FFYCNTSKLF N......... ...STWDNSN STANHTGS.N DTITLQCRIK
CRF02_AG_S  FFYCNTSNLF NRTWNHNGTW NAPGPFNDTE DKTINGTE.D KTITLQCRIK
CRF02_AG_S  FFYCNTAELF NSTWASN... .TNGIWASNI NASNNKDA.N DTITLKCKIK
CRF03_AB_R  FFYCNTTKLF NSTWNGTEE. ......LN.. .NTEG..... DIVTLPCRIK
CRF03_AB_R  FFYCNTTKLF NSTWNNTEE. ......SN.. .NTKG..... DIVTLPCRIK
CRF04_cpx_  FFYCNTTPLF NSTHMQNGT. ......NIT. S.TDSTN... STITLQCRLK
CRF04_cpx_  FFYCNTSGLF NSTYMFNST. ......NRTN T.TNGTN... STITLPCRIK
CRF04_cpx_  FFYCNTSDLF NRTYMVNKN. ......ETNS T.NTTDE... KIIRLPCRIK
CRF05_DF_B  FFYCDTSKLF NATVFNDTV. ......FNAT MFNND...SD KNIILPCKIK
CRF05_DF_B  FFYCNTSGLF NVTVP..... .......... ..NNE..... .TITLPCRIK
CRF06_cpx_  FFYCNTSNLF NTSDLFNTS. ......R..G NDTN...... TTITLPCKIK
CRF06_cpx_  FFYCNTSQLF NNNITDSNE. .......... ...T...... TNFTLPCKIK
CRF06_cpx_  FFYCNTSQLF NSSIPESNE. .......... ...T...... DIITLPCKIK
CRF06_cpx_  FFYCNTSQLF NSSNLNNNS. .......... SDNN...... GTITLPCRIN
CRF11_cpx_  FFYCNTSGLF NNTWLFNST. ......WNSS QELNGT...E PNITLPCRIK
CRF11_cpx_  FFYCNTSGLF NSTWYANDN. ......TSTQ NDMQSN...D .TITLPCRIK
D_CD_84ZR0  FFYCNTSGLF NSAWNISGH. ......STGL N......D.. TIITIPCRIK
D_CD_ELI_K  FFYCNTSGLF NSTWNISAW. ......NNIT ESNNS.TN.. TNITLQCRIK
D_CD_NDK_M  FFYCNTSRLF NSTWNQTNS. ......TGFN .......N.. GTVTLPCRIK
D_UG_94UG1  FFYCNTTRLF NSTWKRNNS. ......EWRS D..NT.PD.. ETITLQCRIK
F1_BE_VI85  FFYCDTSGLF NDTGSN.... .......... .........N GTITLPCRIK
F1_BR_93BR  FFYCNTDELF NDTKFND... .......... ...TG...FN GTITLPCRIK
F1_FI_FIN9  FFYCNTSLLF NNTVPN.... .......... .........N GTITLPCRIK
F1_FR_MP41  FFYCDTSGLF NESEKY.... .......... .........N GTIILPCKIK
F2_CM_MP25  FFYCNTTILF NHTRVNDIL. ......SNNH TR......EN DTITLPCRIK
F2KU_BE_VI  FFYCNTTRLF NDTLNHT... .......... .........ID QNITLPCKIK
G_BE_DRCBL  FFYCNTSGLF NNSILKSNI. .......... SENN...... DTITLNCKIK
G_NG_92NG0  FFYCNTSGLF NNNISNIN.. .......... ...N...... ETITLPCKIK
G_SE_SE616  FFYCNTSGLF NSSLLRSNS. .......... SE.N...... GTITLPCKIK
H_BE_VI991  FFYCNTTKLF NSTWTNSSY. ......TNDT YNSNSTEDIT GNITLQCKIK
H_BE_VI997  FFYCNTSGLF NSSWTGDNI. ......NMPN DTG....... KNITLPCRIK
H_CF_90CF0  FFYCNTSGLF NSSWEMHTN. ......YTSN DTKG...N.. ENITLPCRIK
J_SE_SE702  FFYCNTSTLF NSSWDENNI. ......KDTN STNDN..... TTITIPCKIK
J_SE_SE788  FLYCNTSKLF NSSWDKNSI. ......EATN DTSX...... ATITIPCKIK
K_CD_EQTB1  FSYCDTTDTV DDTEEE.... .......... .........ED TTITIPCRIK
K_CM_MP535  FFYCNTTKLF NETGE..... .......... .........N GTITLPCRIK
N_CM_YBF30  FFYCNTSKLF NEELLN.... ......ETG. .......... EPITLPCRIR
O_CM_ANT70  FFYCNTAKMF NYTFS..... ....CNGTTC SVSNVSQ.G. NNGTLPCKLR
O_CM_MVP51  FFYCNTSGMF NYTFIN.... ....CTKSGC QEIKGSNETN KNGTIPCKLR
O_SN_99SE_  FFYCNTSKMF NYTFS..... ....CIGTNC TSNQNSSNS. NDTRIYCRIK
O_SN_99SE_  FFYCNTSQMF NYTFS..... ....CTRTNC IRQSNSS... INGTISCRIK
U_CD___83C  FFYCNTSELF TGIWNG.... ......TWDK NCTSTESNCT GNITLPCRIK
```

```
            501                                                    550
00BW0762_1  QIINMWQGVG KAMYAPPIAG NIICKSNITG LLLTRDGGEE N......TTE
00BW0768_2  QIINMWQEVG.RAMYAPPIEG NITCKSNITG LLLVRDGGKT ED....NKSE
00BW0874_2  QIINLWQEVG RAIYAPPIAG NITCKSNITG LLLTRD.GG. NNS....TTE
00BW1471_2  QIINMWQGVG QAMYAPPIAG NITCRSNITG LLLTRDGGIN ...EDDNNTE
00BW1616_2  QIINLWQGVG RAMYAPPIAG NITCKSNITG LLLTRDGGGE N.....NSTE
00BW1686_8  QIINMWQEVG RAIYAPPIAG KITCISNITG TLLTRDGGVS NTTE...GNE
00BW1759_3  QIINMWQEVG RAMYAPPIEG NITCNSSITG LLLTRDGGKN S...TNNGTE
00BW1773_2  QIINMWQKVG RAMYAPPIAG NITCKSNITG LLLTRDGGNT S.....STEE
00BW1783_5  QIINMWQGVG QAIYAPPIAG NITCKSNITG LLLTRDGG.. NN...TENTE
00BW1795_6  QIINMWQKVG RAMYAPPIEG NITCISNITG LLLTRDGG.. YE...ANHTE
00BW1811_3  QIINLWQEVG RAMYAPPIAG NITCKSNITG LLLTRDGGGS NTTN...ATE
00BW1859_5  QIINMWQEVG RAMYAPPIAG NITCKSKITG LLLTRDGGKQ ......NESK
00BW1880_2  QIINMWQGVG RAMYAPPIEG NITCNSNITG LLLTRNRGRE NGD...NTTE
00BW1921_1  QIINMWQGVG RAIYAPPIEG NITCKSNITG LLLTRDGGKG NDT.....AE
00BW2036_1  QIINMWQKVG RGIYAPPIEG SITCNSNITG LLLVRDGG.. IN...TSTVE
00BW2063_6  QIINMWQGVG RAMYAPPIAG NITCTSNITG LILTRDGGG. NE...TNETE
00BW2087_2  QIINMWQEVG RAMYAPPIAG NITCKSNITG ILLTRDGGED TKN....KTE
00BW2127_2  QIVNMWQGVG RAIYAPPIAG NITCNSSITG LLLLRDGGTE TENN...RTE
00BW2128_3  QIINLWQEVG RAMYAPPIEG NITCKSNITG LLLTRDGGTN ..N...NNTE
00BW2276_7  QIINMWQGVG RAIYASPIEG SITCKSNITG LLLVHDGG.. NSNT...STE
00BW3819_3  QIINMWQEVG RAIYAPPIAG NITCTSNITG LLLTRDGEPS TE........
00BW3842_8  QVINMWQRVG QAIYAPPIEG IITCNSSITG LLLVRDGD.. NQ...TSDTE
00BW3871_3  QIINMWQEVG RAIYAPPIRG IITCTSNITG LLLTRDGGNT GGN....TTE
00BW3876_9  QIINMWQEVG RAMYAPPIAG NITCTSNITG LLLTRDGG.N GG....NNTE
00BW3886_8  QFIRMWQRVG QAMYAPPIAG NITCRSNITG LLLTRDG... .....KNDTE
00BW3891_6  QIINMWQGVG RAMYAPPIAG RIICKSNITG LLLVRDGGQD N...VMNATE
00BW3970_2  QIINMWQKVG RAIYAPPIAG KITCKSNITG LLLVRDGGGG NN....TATE
00BW5031_1  QIINMWQGVG RAMYAPPIAG NIICKSNITG VLLTYDGGEE N........E
 96BW01B21  QIINMWQGVG RAMYAPPIKG SITCRSNITG LLLTRDGGLN RS...TEEPE
  96BW0407  QIINMWQGVG RAIYAPPIAG NITCVSNITG LLLTWDGGHQ SN.......E
  96BW0502  QIINMWQKVG RAMYAPPIAG NLTCESDITG LLLTRDGGKT G....PNDTE
 96BW06_J4  QIINMWQEVG RAIYAPPIAG NITCKSNITG LLLTRDGGLN NDS......E
 96BW11_06  QFINLWQEVG RAMYAPPIAG NIICKSNITG LLLTRDG... .D...KNDSE
  96BW1210  QIINRWQEVG RAMFAPPIAG NITCKSNITG ILLVRDGGNT SEN.....IE
 96BW15B03  QIINMWQKVG RAIYAPPIEG NITCSSSITG LLLARDGG.. LD...NVTTE
 96BW16_26  QIINMWQGVG RAMYAPPIEG NITCKSNITG LLLVRDGGTE ENN...TGTE
 96BW17A09  QIINMWQGXG QAMYVPPIAG NITCRSNITG LLLTRDGGK. ...VTGNTTE
 96BWMO1_5  QIINMWQGVG RAMYASPIAG NITCKSNITG LLLTRDGG.. NE...TSGIE
 96BWMO3_2  QIINTWQEVG RAIYAPPIAG NIICISNITG LLLTRDGGKT ND...TNDTE
98BWMC12_2  QIINMWQEVG RAMYAPPIAG NITCRSNITG LLLTRD.GGN TTE....TKE
98BWMC13_4  QIINMWQGVG RAMYAPPIAG NITCISNITG LILTRDGG.. VN...RSDTE
98BWMC14_a  QIINMWQEVG RAIYAPPIKG NITCESNITG LLLTRDGGSN DTT......E
98BWMO14_1  QIINMWQGVG QAMYAPPIAG NITCKSNITG ILLTRDGGIN NTN....GTE
98BWMO18_d  QIINMWQKVG RAIYAPPIAG NITCSSRITG LLLTRDGGKN ......DTHE
98BWMO36_a  QIINMWQEVG RAMYAPPIAG NITCKSNITG LLLVRDGGNN NTT......E
98BWMO37_d  QIINMWQKVG RAMYANPIEG NITCRSNITG LLLENDG... N........M
99BW3932_1  QIINMWQKVG RAMYAPPIAG NITCKSNITG LLLVRDGGTA TD.......E
99BW4642_4  QIINMWQEVG RAMYAPPIAG NITCQSNITG LLLTRDGGTE TD....NKTE
99BW4745_8  QIINMWQEVG RAMYAPPIEG NITCKSNITG LLLVRDGGGK N...ATNDTE
99BW4754_7  QIINMWQEVG RAMYAPPIAG RIICNSTITG LILTRDGGNT N......NTE
99BWMC16_8  QIINRWQEVG RAMYAPPIAG NITCTSNITG LLLVRDGGRT SD....STKE
A2_CD_97CD  QIINMWQRVG RAMYAPPIAG VIKCTSNITG MILTRDG..G KNS....INE
A2_CY_94CY  QIINMWQRVG RAMYAPPIAG IIKCTSNITG IILTRDG..G NNG....TNE
A2D___97KR  QIVNMWQRVG RAMYAPPING TIKCTSNITG MILTRDGNSG GNA....TNE
A2G_CD_97C  QIINMWQRVG RAMYAPPIAG IINCTSNITG IILTRDGEKG GDN....TIE
A_BY_97BL0  QIINMWQRVG QAMYAXPIKX SIRCESNITG LLLTRDGXGX TNX....SNE
A_KE_Q23_A  QIINMWQRAG QAMYAPPIPG VIKCESNITG LLLTRDGGKD NN.....VNE
A_SE_SE659  QIINMWQRAG KAMYAPPIQG VIRCESNITG LILTRDG.GD AG.....ENE
A_SE_SE725  QIINMWQRAG QAIYAPPIPG IIRCESNITG LLLTRDG.GV VNS....TNE
A_SE_SE753  QIINMWQRVG QAMYAPPIRG AIRCKSNITG LLLTRDGGNS NSS....TNE
```

```
A_SE_SE853   QIINMWQRAG KAIYAPPIPG IIKCVSNITG LILTRDG.GS NNS....TNE
A_SE_SE889   QIINMWQRAG QAIYAPPIQG VIRCESNITG LILTRDG.GN DNN....ESE
A_SE_UGSE8   QIINMWQRTG QATYAPPIPG VIQCRSNITG LLLTRDGGVT NNT....NNE
A_UG_92UG0   QIINMWQRVG QAMYAPPIQG VIKCESNITG LILTRDG.GV NSS....DSE
A_UG_U455_   QIINMWQRVG QAMYAPPIQG VIRCESNITG LLLTRDG.GT NNT....KNE
AC_IN_2130   QIINMWQRVG QAMYAPPIQG IIKCVSNITG LILTRDGK.S SNS....TDE
AC_RW_92RW   QIINMWQRTG QAMYAPPIQG VISCVSNITG LLLTRDG.GN NNT....TTE
AC_SE_SE94   QIIRMWQRTG QAIYAPPIPG EINCVSNITG LLLTRDG..G NNI....TNE
ACD_SE_SE8   QIINMWQRVG QAMYALPIRG VIRCESNITG LILTRDG.GN NTS....TNE
ACG_BE_VI1   QIINMWQEVG RAMYANPIAG NITCNSNITG LLLTRDGGVN ET....TETE
AD_SE_SE69   QIINMWQRAG RAIYAPPIQG VINCVSDITG LILTRDGGVN .NT.N....E
AD_SE_SE71   QIINMWQRVG QAMYAPPIQG VIKCTSNITG LILTRDG.GG NNS....INA
ADHK_NO_97   QIVNMWQRVG QAMYAPPIKG NITCVSNITG LILTIDXG..  ..N.MSAENF
ADK_CD_MAL   QIINMWQKTG KAMYAPPIAG VINCLSNITG LILTRDGGNS .SD.NS.DNE
AG_BE_VI11   QIVRMWQRVG QAMYAPPIAG KITCRSNITG LILTRDGGNP N....NTNNE
AG_NG_92NG   QIVRMWQRVG QAMYAPPIAG DITCRSNITG LLLTRDGGVN N.....TGNE
AGHU_GA_VI   QIVNMWQRVG RAMYAPPIAG NITCRSNITG IILTRDGG..  .SN.NESTNE
AGU_CD_Z32   QIVNMWQRVG QAMYAPPIKG VIKCESNITG ILLTRDGVG. NNT....ANE
AJ_BW_BW21   QIVRMWQRVG QAIYAPPIAG NITCTSNITG LLLTRDGGYT ..N.NTNGTE
B_AU_VH_AF   QIINMWQKVG KAMYAPPING QIRCSSNITG LILTRDGGNQ .....ENKTE
B_CN_RL42_   QIVNMWQEVG KAMYAPPIEG QIRCSSNITG LLLTRDGGNN .E..S.KPTE
B_DE_D31_U   QIINMWQEVG KAMYAPPISG QIRCSSNITG LLLTRDGGKN .K..D.NETE
B_DE_HAN_U   QIINMWQEVG KAMYAPPIGG LIRCSSNITG LILTRDGGND .N..S.STTE
B_FR_HXB2_   QIINMWQKVG KAMYAPPISG QIRCSSNITG LLLTRDGGNS .N..N..ESE
B_GA_OYI__   QIVNMWQEVG KAMYAPPISG QIRCSSKITG LLLTRDGGKN ....TTNGIE
B_GB_CAM1_   QIINRWQEVG KAMYAPPITG TISCSSNITG LLLTRDGGRG .E..N..ETE
B_GB_GB8_C   QIVNMWQEVG KAMYAPPITG QIRCASHITG LLLTRDGGRE .N..NTNETE
B_GB_MANC_   QILNLWQEVG KAMYAPPISG QISCSSNITG LLLTRDGGNT .NT.TGNTTE
B_KR_WK_AF   QIINRWQEVG KAMYAPPISG LIRCSSNITG LLLTRDGGNE .NN.GTNGTE
B_NL_3202A   QIINMWQGVG KAMYAPPISG QIRCSSNITG LLLTRDGGKD .E..NKTGTE
B_TW_TWCYS   QIINMWQRVG KAMYAPPIEG LIKCSSNITG LMLTRDGGTN .D....SEVE
B_US_BC_L0   QIINMWQEVG KAMYAPPISG QIRCTSNITG LLLTRDGGTS .D..T.NTTE
B_US_DH123   QIINMWQEVG KAMYAPPISG QIWCSSNITG LLLTRDGGKN ......SSTE
B_US_JRCSF   QIINMWQEVG KAMYAPPIKG QIRCSSNITG LLLTRDGGK. ....NESEIE
B_US_MNCG_   QIINMWQEVG KAMYAPPIEG QIRCSSNITG LLLTRDGGKD .T..DTNDTE
B_US_P896_   QIINMWQKVG KAMYAPPITG QIRCSSNITG LLLTRDGGNS ....TETETE
B_US_RF_M1   QIVNMWQEVG KAMYAPPISG QIKCISNITG LLLTRDGGED .T..T.NTTE
B_US_SF2_K   QIINMWQEVG KAMYAPPIGG QISCSSNITG LLLTRDGGTN .V..T.NDTE
B_US_WEAU1   QIINRWQEVG KAMYAPPIEG QIRCLSNITG LLLTRDGGSS .E..E.NQTE
B_US_WR27_   QIINMWQEVG KAMYAPPIDG QIRCSSNITG LLLTRDGGN. ....SNETTE
B_US_YU2_M   QIINMWQEVG KAMYAPPIRG QIRCSSNITG LLLTRDGGK. ....DTNGTE
BF1_BR_93B   QIVNMWQEVG RAMYAAPIAG NITCSSNITG LLLTRDGG.. Q...NNQTEE
C_BR_92BR0   QIINMWQGVG RAMYAPPIEG ILTCRSNITG LLLTRDGGTG .....MHDTE
C_BW_96BW0   QIINMWQGVG RAIYAPPIAG NITCISNITG LLLTRDGGTT RNN....ESE
C_BW_96BW1   QFINLWQEVG RAMYAPPIAG NIICKSNITG LLLTRDG... .D...KNDSE
C_BW_96BW1   QIINRWQEVG RAMFAPPIAG NITCKSNITG ILLVRDGGNT SEN.....IE
C_BW_96BW1   QIINMWQKVG RAIYAPPIEG NITCSSSITG LLLARDGG.. LD...NVTTE
C_ET_ETH22   QIINMWQGVG RAMYAPPIEG IIMCRSNITG LLLTRDGAKE PH....STKE
C_IN_93IN1   QIINMWQEVG RAMYAPPIAG NITCTSNITG LLLVHDGGIK EN.DTENKTE
C_IN_93IN9   QIINMWQEVG RAMYAPPIEG NITCKSNITG LLLVRDGGAE AK...TNNTE
C_IN_93IN9   QIINMWQEVG RAMYAPPIAG NITCKSNITG LLLVRDGGRG ND..TENNTE
C_IN_94IN1   QIINMWQEVG RAMYAPPIAG NITCKSNITG ILLERDG..G SG...SNGTE
C_IN_95IN2   QIINMWQEVG RAMYAPPIEG KITCRSNITG LLLVRDGGED KNNTETNKTE
CRF01_AE_C   QVVNMWQEVG KAMYAPPISE AVNCVSNITG IILTRDGGNA TNET......
CRF01_AE_C   QIVNMWQEVG RAMYAPPISE VINCVSNITG ILLTRDGGIN QNQTNK..NE
CRF01_AE_C   QVIKMWQEVG QAMYAPPIDE AINCVSNITG ILLVRDGGKI ENET....IE
CRF01_AE_T   QIINMWQGTG QAMYAPPIDG KINCVSNITG ILLTRDGG.. ANNTS...NE
CRF01_AE_T   QIIRMWQGAG QAMYAPPISG IINCVSNITG ILLTRDGGS. ANNTN...NE
CRF01_AE_T   QIINMWQGVG QAMYNPPISG NINCVSNITG ILLTRDGGGG NGTNN...EE
CRF01_AE_T   QIINMWQEVG QAMYAPPITG KINCVSNITG ILLTRDGG.. ANNKS...SE
CRF01_AE_T   QIINMWQGAG QAMYAPPISG RINCVSNITG ILLTRDGG.. VNNTD...NE
```

```
CRF01_AE_T  QIINMWQGAG QAMYAPPING TINCISNITG ILLTRDGGD. NNNTI...NE
CRF02_AG_F  QIVNMWQKVG LAMYAPPISG EIRCKSNITG LLLTRDG.GS NNS....TNE
CRF02_AG_F  QIVNMWQKVG RAMYAPPIPG EIRCESNITG LLLTRDG.GS NNS....TNE
CRF02_AG_G  QIINMWQKVG QAMYAPPIQG VIRCDSNITG LLLTRDG.GS NNN....TPE
CRF02_AG_N  QIVNMWQKVG QAMYAPPIQG IIRCDSNITG LLLTRDG.G. NNS....TNE
CRF02_AG_S  QIVRMWQKVG QAMYAPPIPG EIRCESNITG LLLTRDG.GN DNN....NTE
CRF02_AG_S  QIINMWQKVG QAIYAPPIEG VIRCDSNITG ILLTRDG.GD NTN....GDE
CRF03_AB_R  QIINMWQEVG KARYAPPIAG QIRCSSNITG LLLTRDGGNQ .S....NVTE
CRF03_AB_R  QIINMWQEVG KAMYAPPIAG QIRCSSNITG LLLTRDGGNQ .N....NVTE
CRF04_cpx_  QFVRMWQEVG QAMYASPIAG SINCSSDITG IILTRDG...  ....GTNNTE
CRF04_cpx_  QIVRMWQGVG QAMYAPPIAG SINCSSDITG IILTRDGGIS NNN.ETNDNE
CRF04_cpx_  QIVNRWQEVG QAIYAPPLQG SLTATQVITG IILTRDGG..  .NR.SDTGNE
CRF05_DF_B  QIVRMWQGVG QAMYAAPIAG NIACNSTITG ILLARDGGNG .ND.SSNDTE
CRF05_DF_B  QIINMWQGVG QAMYAAPIAG NITCNSNITG ILLTRDG..G .VN.ITNDTE
CRF06_cpx_  QIVRMWQRVG QAMYAPPIAG NITCVSNITG IILTRDGN.N EN.....VSE
CRF06_cpx_  QIVRMWQRVA QAMYAPPIAG NIICTSNITG LLLTRDGGRN DS.....NSE
CRF06_cpx_  QIVRMWQRVG QAIYAPPIAG NITCISNITG LLLTRDGN.T NT.....TSE
CRF06_cpx_  QIIRMWQRVG QAMYAPPIAG NITCTSNITG LLLTRDGH.N D.......TE
CRF11_cpx_  QIVRMWQRVG QAMYAPPIQG EIRCDSNITG LLLTRDGG..  ....LNSTNE
CRF11_cpx_  QIINMWQRVG QAVYAPPIQG ELRCDSNITG LLLTRDGGEG ..N.DTIGKE
D_CD_84ZR0  QIINMWQEVG KAMYAPPIEG QINCSSNITG LLLTRDGGAN .NT.Q...ND
D_CD_ELI_K  QIIKMWQAG. .AIYAPPIER NILCSSNITG LLLTRDGGIN .NS.T...NE
D_CD_NDK_M  QIVNLWQRVG KAMYAPPIEG LIKCSSNITG LLLTRDGGAN .NS.S...HE
D_UG_94UG1  QIINMWQEVG KAMYAPPIEG FINCSSNITG LLLTRDGGAI .NS.SQ..NE
F1_BE_VI85  QIVNMWQGVG RAMYTSPIAG NITCNSNITG LLLTRDGG..  ....NESNIE
F1_BR_93BR  QIVNMWQEVG RAMYANPIAG NITCNSNITG LLLTRDGG..  ....LNSTNE
F1_FI_FIN9  QFVNMWQEVG RAMYAAPIAG NITCNSNITG LLLTRDGG..  QS..NNSDSE
F1_FR_MP41  QIINMWQGVG QAMYSAPIAG RINCNSTITG LLLTRDGG..  QSN.DTNRTE
F2_CM_MP25  QIVNMWQRVG QAMYAPPIAG KIQCNSNITG LLLTIDGG..  ....EGNESE
F2KU_BE_VI  QIINRWQGVG QAMYAPPIAG NITCRSNITG MILTRDGGNS N...DTIDNE
G_BE_DRCBL  QIVRMWQRVG QAMYAPPIAG NITCRSNITG LILTRDGGDN N.....STSE
G_NG_92NG0  QIVRMWQKVG QAMYALPIAG NLVCKSNITG LILTRDGGNN N....DSTEE
G_SE_SE616  QIVRMWQRVG QAMYAPPIAG NIECNSSITG LILTRDGGNN NNT.NTSESE
H_BE_VI991  QIVNMWQRVG QAMYAPPIRG NITCISNITG LILTFD....  ..R.NNTNNV
H_BE_VI997  QIVNMWQRVG QAMYAPPIKG SITCVSNITG LILTYDED..  ..K.GNNDNV
H_CF_90CF0  QIVNMWQRVG RAMYAPPIQG NIMCVSNITG LILTIDEG..  ..N.ASAENY
J_SE_SE702  QIVRMWQRTG QAIYAPPIAG NITCKSNITG LLLTRDGGNR .NG.SENGTE
J_SE_SE788  QIVRMWQRTG QAIYAPPIAG NITCTSNITG LLLTRDGGNR GNG.SENGTE
K_CD_EQTB1  QIINMWQKVG QAIYAPPTAG NITCRSNITG MILTRDGGND N...NTRTEE
K_CM_MP535  QIINMWQKVG KAIYAPPIAG SINCSSNITG MILTRDGGNN ......THNE
N_CM_YBF30  QIVNLWTRVG KGIYAPPIRG VLNCTSNITG LVLEYSGGPD ......TKET
O_CM_ANT70  QVVRSWIRGQ SGLYAPPIKG NLTCMSNITG MILQMDNTWN SSNN....NV
O_CM_MVP51  QLVRSWMKGE SRIYAPPIPG NLTCHSNITG MILQLDQPWN STGE.....N
O_SN_99SE_  QVVRSWIQGG SGLYAPPRKG NLTCSSLITG MILQLDMPWN STNNS...NA
O_SN_99SE_  QVVRSWIQGG SGLYAPPRPG YLTCNSSITG MILQLDKTWN RTNNS...ES
U_CD___83C  QVVRTWQGVG QAMYAPPIEG TIRCSSNITG LLLTRDGGNG N....ATQNE
```

```
                551                                                      600
00BW0762_1  TFRPAGGDMR DNWRSELYKY KVVEIKPLGI APTSAKRRVV EREKR.....
00BW0768_2  IFRPGGGDMR DNWRSELYKY KVVEIKPLGV APTEAKRRVV EREK.R....
00BW0874_2  IFRPQGGNMK DNWRSELYKY KVVEVKPLGV APTKAKRRVV EREKR.....
00BW1471_2  IFRPGGGNMR DNWRSKLYKY KVVEIKPLGV APNKAKRRVV EREK.R....
00BW1616_2  TFRPAGGEMR DNWRSELYKY KVVEVKPLGI APTEAKRRVV QREKR.....
00BW1686_8  TFRPGGGDMR NNWRSELYKY KVVEIRPLGV APTEARRRVV EREK.R....
00BW1759_3  IFRPGGGDMR DNWRSELYKY KVVKIKPLGI APTKAQRRVV KREKR.....
00BW1773_2  IFRPEGGDMR DNWRSELYKY KVVEIKPLGV APTKAKRRVV EREK.R....
00BW1783_5  TFRPGGGDMR DNWRNELYKY KVVEIKPLGI APTSAKRRVV EREK.R....
00BW1795_6  IFRPIGGDMR DNWRSELYKY KVVEIKPLGL APTESKRRVV EREK.R....
00BW1811_3  TFRPGGGDMR DNWRSELYKY KVVEVKPLGL APTEAKRRVV EREK.R....
00BW1859_5  IFRPGGGDMR NNWRSELYKY KVVEIKPLGL APTGAKRRVV EREK.R....
00BW1880_2  TFRPAGGDMR DNWRSELYKY KVVEIKPLGI APTKAKRRVV EREKR.....
00BW1921_1  IFRPEGGDMK NNWRSELYKY KVVEIRPLGV APTKAKRRVV EREK......
00BW2036_1  TFRPEGGNMR DNWRSELYKY KVVEIKPLGV APTEAQRRVV EKQK.R....
00BW2063_6  TFRPAGGDMR DNWRSELYKY KVVEIKPLGL APTKAKRRVV EREK.R....
00BW2087_2  TFRPGGGDMR DNWRSELYKY KVVEIKPLGV APTKAKRRVV EREK.R....
00BW2127_2  TFRPGGGDMR DNWRSELYKY KVVEIKPLGV APTKAKRRVV EREK.R....
00BW2128_3  TFRPVGGDMR DNWRSELYKY KVVEIKPLGV APTEAKRRVV KREK.R....
00BW2276_7  IFRPGGGDMR DNWRSELYKY KVVEVKPLGI APTEAKRRVV EREK.R....
00BW3819_3  TFRPGGGDMR DNWRSELYKY KVVEVKPLGI APTGAKRRVV EREK.R....
00BW3842_8  TFRPQGGEMR DNWRSELYKY KVVEIKPLGV APTTAKRRVV EREK.R....
00BW3871_3  IFRPEGGDMR NNWRNELYKY KVVEIKPLGI APTGAKRRVV EREK.R....
00BW3876_9  TFRPGGGNMK DNWRSELYKY KVVEIKPLGV APTEAKRRVV EREK.R....
00BW3886_8  TFRPGGGNMR DNWRNELYKY KVVEIKPLGI APTEAKRRAV EREK.R....
00BW3891_6  TFRPGGGDMR DNWRSELYKY KVVEIKPLGV APTSAKRRVV EREK......
00BW3970_2  IFRPGGGNMK DNWRSELYKY KVVEIKPLGI APTGAKRRVV GREK.R....
00BW5031_1  TFRPAGGNMK DNWRSELYKY KVVEIKPLGI APTKAKRRVV EREK......
 96BW01B21  IFRPGGGDMR NNWRSELYKY KVVEIKPLGV APTGAKRRVV EREK.R....
  96BW0407  TFRPGGGDMR DNWRSELYKY KVVEIKPLGI APTEARRRVV EREKR.....
  96BW0502  IFRPGGGDMR DNWRNELYKY KVVEIKPLGV APTEAKRRVV EREK.R....
 96BW06_J4  TFRPIGGEMR NNWRSELYKY KVVEIKPLGI APTKAKRRVV EREK.R....
 96BW11_06  TFRPAGGDMR DNWRSELYKY KVVEIKPLGL APTKAKRRVV EREK.R....
  96BW1210  TFRPGGGNMK DNWRSELYKY KVVEVKPLGI APTRAKRRVV EREK.R....
 96BW15B03  IFRPQGGDMK DNWRNELYKY KVVEIKPLGV APTEAKRRVV EREK.R....
 96BW16_26  IFRPEGGDMR DNWRSELYKY KVVEIKPLGI APTKAKRRVV EREK.R....
 96BW17A09  TFRPGGGNMR DNWRSELYKY KVVEVKPLGV APTAAKRRVV EREK.R....
 96BWMO1_5  IFRPAGGDMR DNWRSELYKY KVVEIKPLGL APTKSKRRVV GREK.R....
 96BWMO3_2  IFRPGGGNMK DNWRSELYKY KVVEIKPLGV APTKAKRRVV EREK.R....
98BWMC12_2  TFRPGGGDMR DNWRSELYKY KVVEIKPLGV APTEAKRRVV ERGKR.....
98BWMC13_4  IFRPAGGDMR DNWRSELYKY KVVEIKPLGL APTKAKRRVV EREK.R....
98BWMC14_a  TFRPEGGDMR NNWRSELYKY KVVEIRPLGI APTGAKRRVV NREK.R....
98BWMO14_1  TFRPGGGDMR DNWRSELYKY KVVEVKPLGI APTKAQRRVV EREK.R....
98BWMO18_d  TFRPAGGDMR DNWRSELYKY KVVEIKPLGV APSEAKRRVV EREK......
98BWMO36_a  TFRPGGGNMK DNWRSELYKY RVVEIKPLGI APTGAKRRVV EREK.R....
98BWMO37_d  TFRPGGGDMK DNWRSELYKY KVVEIKPLGI APTEAKRRVV EREK.R....
99BW3932_1  IFRPGGGDMR DNWRSELYKY KVVEIKPLGI APTEAKRRVV EREK.R....
99BW4642_4  TFRPGGGDMR DSWRSELYKY KVVEIKPLGV APTKAKRRVV EREK.R....
99BW4745_8  IFRPEGGDMR NNWRSELYKY KVVEIKPLGV APTKAKRRVV DKEK......
99BW4754_7  IFRPIGGNMR DNWRSELYKY KVVEIKPLGI APTKAKRRVV EREKR.....
99BWMC16_8  IFRPGGGDMR DNWRSELYKY KVVEIKPLGI APTEAKRRVV EREKR.....
A2_CD_97CD  TFRPGGGDMR DNWRSELYKY KVVKIEPLGI APTEARRRVV QREK.R....
A2_CY_94CY  TFRPGGGDMR DNWRSELYKY KVVKLEPLGV APTRAKRRVV EREK.R....
A2D___97KR  TFRPGGGDMR DNWRSELYKY KVVKLEPLGV APTRARRRVV EREK......
A2G_CD_97C  VFRPVGGDMR DNWRSELYKY KVVKIKPLGI APTRARRRVV EKEK.R....
A_BY_97BL0  TFRPIXGDXR NNWRSELYKX KVVKIEPIXV APTRAKRRXX EREK.R....
A_KE_Q23_A  TFRPGGGDMR DNWRSELYKY KVVEIEPLGV APTRAKRRVV EREK.R....
A_SE_SE659  TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTQARRRVV KREK.R....
A_SE_SE725  TFRPGGGNMK DNWRSELYKY KVVKIEPLGV APTRARRRVV QREK.R....
A_SE_SE753  TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV EREK.R....
```

```
A_SE_SE853   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV EREK.R....
A_SE_SE889   IFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV EREK.R....
A_SE_UGSE8   TFRPGGGDMR DNWRSELYKY KVVKLEPLGV APTKAQRRVV KREK.R....
A_UG_92UG0   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKARRRVV EREK.R....
A_UG_U455_   TFRPGGGDMR DNWKSELYKY KVVKIEPLGV APTRAKRRVV EREK.R....
AC_IN_2130   IFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTRARRRAV GREK.R....
AC_RW_92RW   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTRAKRRVV EREK.R....
AC_SE_SE94   TFRPGGGDMQ DNWRSELYKY KVVQIEPLGV APTKARRRVV EREK.R....
ACD_SE_SE8   TIRPAGGDMR DNWRSELYKY KVVKIEPLGV APTKARRRVV EREK.R....
ACG_BE_VI1   IFRPGGGNMK DNWRSELYKY KVVEIKPLGV APTKAKRRVV EREK.R....
AD_SE_SE69   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTRAKRRVV EREK.R...V
AD_SE_SE71   TFRPGGGDMR DNWRSELYKY KVVKIEPVGI APNRAKRRVV EREK.R....
ADHK_NO_97   TFRPGGGDMR DNWRSELYKY KVVXXXPLGV APTXARRRVV QREK.R....
ADK_CD_MAL   TLRPGGGDMR DNWISELYKY KVVRIEPLGV APTKAKRRVV EREK.R...A
AG_BE_VI11   TFRPGGGDMR DNWRSELYQY KVVKIKSLGV APTKARRRVV EREK.R...A
AG_NG_92NG   TFRPGGGDMR DNWRSELYKY KIVKIKPLGI APTKARRRVV ERGK.R...A
AGHU_GA_VI   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTRARRRVV EREK.R...A
AGU_CD_Z32   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV AREK.R....
AJ_BW_BW21   IFTPTGRNMR DNWRSELYKY KVVKIEPIGV APTRAKRRVV GREK.R...A
B_AU_VH_AF   IFRPGGGDMR DNWRSELYKY KVVRIEPLGV APTKAKRRVV QREK.R...A
B_CN_RL42_   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKARRRVV QREK.R...A
B_DE_D31_U   TFRPGGGNMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV QREK.R...A
B_DE_HAN_U   IFRPGGGNMR DNWRNELYKY KVVKIEPLGV APTKAKRRVV QREK.R...A
B_FR_HXB2_   IFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV QREK.R...A
B_GA_OYI__   IFRPAGGDMR DNWRSELYKY KVVKIEPLGV APTKARRRVV QREK.R...A
B_GB_CAM1_   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV QREK.R...A
B_GB_GB8_C   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV QREK.R...A
B_GB_MANC_   TFRPGGGNMR DNWRSELYKY KVVKVEPLGI APTKAKRRVV QREK.R...A
B_KR_WK_AF   TFRPEGGNMK DNWRSKLYKY KVVRIEPLGI APTRARRRVV QREK.R...A
B_NL_3202A   IFRPGGGDMK DNWRSELYKY KVVKIEPLGV APTRAKRRVV QREK.R...A
B_TW_TWCYS   VFRPGGGDMK DIWRNELYKY KVVKVEPLGL APTRARRRVV QREK.R...A
B_US_BC_L0   TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTTAKRRVV QREK.R...A
B_US_DH123   IFRPGGGDMR DNWRSELYKY KVVRVEPLGI APTKAKRRVV QREK.R...A
B_US_JRCSF   IFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV QREK.R....
B_US_MNCG_   IFRPGGGDMR DNWRSELYKY KVVTIEPLGV APTKAKRRVV QREK......
B_US_P896_   IFRPGGGDMR DNWRSELYKY KVVRIEPIGV APTRAKRRTV QREK.R....
B_US_RF_M1   IFRLGGGNMR DNWRSELYKY KVVRIEPLGV APTRAKRRVV QREK.R...A
B_US_SF2_K   VFRPGGGDMR DNWRSELYKY KVIKIEPLGI APTKAKRRVV QREK.R...A
B_US_WEAU1   IFRPGGGNMK DNWRSELYKY KVVKIEPLGV APTKAKRRVV QREK.R...A
B_US_WR27_   IFRPGGGDMR DNWRSXLYKY KVVXIEPLGV APTKXKRRVX XREK.R...X
B_US_YU2_M   IFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV QREK.R...A
BF1_BR_93B   TFRPGGGNMK DNWRSELYKY KVVEIEPLGV APTKAKRQVV KREK.R...A
C_BR_92BR0   IFRPEGGDMR DNWRSELYKY KVVEIKPLGI APTKAKRRVV EREK.R....
C_BW_96BW0   IFGPGGGDMR DNWRSELYKY KVVEIKPLGI APTEARRRVV EREKR.....
C_BW_96BW1   TFRPAGGDMR DNWRSELYKY KVVEIKPLGL APTKAKRRVV EREK.R....
C_BW_96BW1   TFRPGGGNMK DNWRSELYKY KVVEVKPLGI APTRAKRRVV EREK.R....
C_BW_96BW1   IFRPQGGDMK DNWRNELYKY KVVEIKPLGV APTEAKRRVV EREK.R....
C_ET_ETH22   IFRPEGGDMR DNWRSELYKY KVVEIKPLGV APTKPKRRVV EREK......
C_IN_93IN1   IFRPGGGDMR DNWRSELYKY KVVEIKPLGV APTAAKRRVV EREK.R....
C_IN_93IN9   TFRPGGGDMR DNWRSELYKY KVVEIKPLGV APTTAKRRVV EREK.R....
C_IN_93IN9   IFRPGGGDMR NNWRSELYKY KVVEIKPLGV APTKAKRRVV EREKRA....
C_IN_94IN1   TFRPGGGDMR NNWRSELYKY KVVEIQPLGV APTEAKRRVV ERGK.R....
C_IN_95IN2   TFRPGGGDMR DNWRSELYKY KVVEVKPLGV APTTAKRRVV EREK.R....
CRF01_AE_C   .FRPGGGNMK DNWRSELYKY KVVQIEPLGI APTRARRRVV EREK.R....
CRF01_AE_C   TFRPGGGNIK DNWRSELYKY KVVQIEPLGI APTKARRRVV EREK.R....
CRF01_AE_C   TFRPGGGNMK DNWRSELYKY KVVQIEPLGV APTGAKRRVV EREK.R....
CRF01_AE_T   TFRPGGGNIK DNWRSELYKY KVVQIEPLGI APTRAKRRVV EREK.R....
CRF01_AE_T   TFRPEGGNIK DNWRSELYKY KVVQIEPLGI APTRAKRRVV EREK.R....
CRF01_AE_T   TFRPGGGNMK DNWRNELYKY KVVEIEPLGI APTKAKRRVV EREK.R....
CRF01_AE_T   TFRPGGGNIK DNWRSELYKY KVVEIEPLGI APTRAKRRVV EREK.R....
CRF01_AE_T   TFRPGGGNIK DNWRSELYKY KVVQIEPLGI APTRAKRRVV EREK.R....
```

```
CRF01_AE_T  TFRPGGGNIK DNWRSELYKY KVVQIEPLGI APSKAKRRVV EREK.R....
CRF02_AG_F  TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTRPKRRVV EREK.R....
CRF02_AG_F  TFRPGGGDMR DNWRSELYKY KVVKIEPLGI APTHAKRRVV EREK.R....
CRF02_AG_G  IFRPGGGNMR DNWRSELYKY KVVKIEPLGV APTRAKRRVV EREK.R....
CRF02_AG_N  TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTRAKRRVV EREK.R....
CRF02_AG_S  TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APSHAKRRVV EREKRA....
CRF02_AG_S  TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APSHAKRRVV EREK.R....
CRF03_AB_R  IFRTGGGDMR DNWRSELYKY KVVKIEPLGV APTRAKRRVV QREK.R...A
CRF03_AB_R  IFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV QREK.R...A
CRF04_cpx_  IFRPGGGDMR DNWRSELYKY KVVKIEPIGV APNKARRRVV QREK......
CRF04_cpx_  TLRPGGGDMR DNWRSVLYKY KVVKIEPVGI APTGARRRVV QKEK.R....
CRF04_cpx_  TFRPGGGDMR DNWRSELYKY KVVQIEPVGV APTRARRRVV QREK.R....
CRF05_DF_B  IFRPGGGDMR DNWRSELYKY KVVEIQPLGI APTRAKRQVV KREK.R...A
CRF05_DF_B  TFRPGGGDMR DNWRSELYKY KVVEIEPLGV APTRAKRQVV QREK.R...A
CRF06_cpx_  TFRPGGGDMR DNWRSELYKY KVVKIKPLGI APTWARRRVV GREK.R...A
CRF06_cpx_  TFRPGGGDMR DNWRNELDKY KVVKIKPLGI APTRARRRVV GKEKRA...V
CRF06_cpx_  IIRPGGGDMR DNWRNELYKY KVVKIKPLGI APTEARRRVV GREK.R...A
CRF06_cpx_  IFRPGGGNMK DNWRSELYKY KVVKIKPLGI APTKARRRVV GREKRA...V
CRF11_cpx_  TFRPTGGDMR DNWRSELYKY KVVEIKPLGV APTRAKRRVV EREK.R...A
CRF11_cpx_  TFRPTGGDMR NNWRSELYKY KVVEIKPLGV APTKAKRRVV EREK.R...A
D_CD_84ZR0  TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRRVV EREK.R...A
D_CD_ELI_K  TFRPGGGDMR DNWRSELYKY KVVQIEPLGV APTRAKRRVV EREK.R...A
D_CD_NDK_M  TIRPGGGDMR DNWRSELYKY KVVKIEPIGV APTKARRRVV EREK.R...A
D_UG_94UG1  TFRPGGGDMR NNWRSELYKY KVVKLEPIGL APTAAKRRVV EREK.R...A
F1_BE_VI85  TFRPEGGNMK DNWRSELYKY KVVEIEPLGV APTKAKRQVV QREK.R...A
F1_BR_93BR  TFRPGGGNMK DNWRSELYKY KVVEIEPLGV APTKAKRQVV KRER.R...A
F1_FI_FIN9  TFRPGGGDMK DNWRSELYKY KVVEIEPLGV APTRPKRPVV RRER.R...A
F1_FR_MP41  TFRPEGGNMK DNWRNELYKY KVVEIEPLGV APTKARRRVV QRER.R...A
F2_CM_MP25  TLRPGGGDMR DNWRSELYKY KVVKIEPLGV APTKAKRQVV QREK.R...A
F2KU_BE_VI  TFRPGGGNMK DNWRSELYKY KVVKIEPLGI APTKARRRVV QREK.R...G
G_BE_DRCBL  IFRPGGGDMK NNWRSELYKY KTVKIKSLGI APTRARRRVV EREK.R...A
G_NG_92NG0  TFRPGGGDMR DNWRSELYKY KTVKIKSLGV APTRARRRVV EREK.R...A
G_SE_SE616  IFRPGGGDMR DNWRSELYKY KTVKIKSLGV APTRARRRVV EREK.R...A
H_BE_VI991  TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTEARRRVV EREK.R....
H_BE_VI997  TFRPGGGDMR DNWRSELYKY KVVKIEPLGV APTEARRRVV EREK.R....
H_CF_90CF0  TFRPGGGDMR DNWRSELYKY KVVKIEPLGI APTKTRRRVV EREK.R....
J_SE_SE702  TFRPTGGNMK DNWRSELYKY KVVELEPLGV APTKAKRRVV EREK.R...A
J_SE_SE788  TFRPTGGNMK DNWRSELYKY KVVEIEPLGV APTKAKRRVV EREK.R...A
K_CD_EQTB1  TFRPGGGDMR DNWRSELYKY KVVQIEPLGI APTRARRRVV QREK.R...A
K_CM_MP535  TFRPGGGDMR DNWRSELYKY KVVQIEPLGI APTRARRRVV QREK.R...A
N_CM_YBF30  IVYPSGGNMV NLWRQELYKY KVVSIEPIGV APGKAKRRTV SREK.R...A
O_CM_ANT70  TFRPIGGDMK DIWRTELFNY KVVRVKPFSV APTRIARPVI STRTHR.EKR
O_CM_MVP51  TLRPVGGDMK DIWRTKLYNY KVVQIKPFSV APTKMSRPII NIHTPHREKR
O_SN_99SE_  TFRPTGGDMK DIWRTELFKY KVVKVKPFSV APTKIARPVI GTGTQR.EKR
O_SN_99SE_  TFRPIGGDMK DIWRTELFKY KVVKIKPFSV APTKIARPVI GTGTRR.EKR
U_CD___83C  TFRPGGGDMK DNWRSELYKY KVVKIEPLGV APTRAKRRVV EREK.R....
```

```
            601                                                       650
00BW0762_1  AVGIGAVFLG .FLGAAGSTM GAASITLMVQ ARQLLSGIVQ QQNNLLRAIE
00BW0768_2  AVGIGAVLLG .FLGAAGSTM GAASITLTVQ ARQVLSGIVQ QQSNLLRAIE
00BW0874_2  AVGIGAVFLG .FLGAAGSTM GAASMTLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW1471_2  AVGLGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLMRAIE
00BW1616_2  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
00BW1686_8  AVGIGAVLLG .FLGAAGSTM GAASMTLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW1759_3  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW1773_2  AVGIGAMFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW1783_5  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ VRQLLSGIVQ QQNNLLRAIE
00BW1795_6  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW1811_3  AVGIGAVFLG .FLGAAGSTM GAASMTLTVQ ARQLLSGIVQ QQSNLLKAIE
00BW1859_5  AVGIGAVFLG .FLGAAGSTM GAASTTLTAQ ARQVLSGIVQ QQSNLLRAIE
00BW1880_2  AVGIGAVFLG .FLGAAGSTM GAASMTLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW1921_1  RAALGAVLLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW2036_1  AVGMGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLKAIE
00BW2063_6  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW2087_2  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW2127_2  AVGMGAVILG .FLGAAGSTM GAASITLTVQ TKQLLSGIVQ QQSNLLKAIE
00BW2128_3  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLFGIVQ QQNNLLRAIK
00BW2276_7  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW3819_3  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLFGIVQ QQSNLLRAIE
00BW3842_8  AVGMGAMILG .FLSAAGSTM GAASITLTIQ ARQLLSGIVQ QQSNLLRAIE
00BW3871_3  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW3876_9  AVEIGAVFLG .FLGAAGSTM GAASITLTAQ ARQLLSGIVQ QQSNLLRAIE
00BW3886_8  AVGIGAVILG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
00BW3891_6  RAAIGAMFLG .FLGAAGSTM GAASITLTAQ ARQLLSGIVQ QQSNLLRAIE
00BW3970_2  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARKLLSGIVQ QQSNLLRAIE
00BW5031_1  RAALGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLKAIE
 96BW01B21  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
  96BW0407  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
  96BW0502  AVGIGAVCLG .FLGAAGSTM GAASITLTVQ ARLLLSGIVQ QQNNLLRAIE
 96BW06_J4  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ AKQLLSGIVQ QQSNLLKAIE
 96BW11_06  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
  96BW1210  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
 96BW15B03  AVGIGAVIFG .FLGAAGSTM GAASITLTAQ ARQLLSGIVQ QQSNLLRAIE
 96BW16_26  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLQAIE
 96BW17A09  AVGLGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
 96BWMO1_5  AVTFGAMFLG .FLGAAGSTM GAASMTLTVQ ARQLLSGIVQ QQSNLLRAIE
 96BWMO3_2  AVGIGAVLLG .FLGTAGSTM GAASITLTVQ ARQVLSGIVQ QQSNLLRAIE
98BWMC12_2  AAGLGAVLFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
98BWMC13_4  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLRAIE
98BWMC14_a  AVGVAAVFLG .FLSAAGSTM GAASITLTVQ ARQSLSGIVQ QQSNLLRAIE
98BWMO14_1  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
98BWMO18_d  RAALGAVFLG .FLGAAGSTM GAASITLTVQ TRKLLSGIVQ QQSNLLKAIE
98BWMO36_a  AVTLGAMFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
98BWMO37_d  AVGIGAVFLG .FLGAAGSTM GAASITLMVQ ARQLLSGIVQ QQSNLLRAIE
99BW3932_1  AVGLGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
99BW4642_4  AVGIGAVLLG .FLGAAGSTM GAASIALTAQ ARQVLSGIVQ QQSNLLRAIE
99BW4745_8  RAIAGAVFLG .FLGVAGSTM GAASVALTVQ ARQLLSGIVQ QQSNLLRAIE
99BW4754_7  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
99BWMC16_8  AVTIGAMFLG .FLSAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
A2_CD_97CD  AVGLGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLTGIVQ QQSNLLKAIE
A2_CY_94CY  AVGLGAVFLG .FLGAAGSTM GAASLTLTVQ ARQLLSGIVQ QQSNLLQAIE
A2D___97KR  RAAVGLFFLG .FLGAAGSTM GAASVTLTVQ ARQLLSGIVQ QQSNLLKAIE
A2G_CD_97C  AVGLGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLKAIE
A_BY_97BL0  AIGLXAAFLX .FLGAAXSTX GAASMTLTVQ ARQLLSGIVQ QQSNLLXAIX
A_KE_Q23_A  AVGIGAVFLG .FLGAAGSTM GATSITLTVQ ARQLLSGIVQ QQNNLLRAIE
A_SE_SE659  AVGLGAVFIG .FLGAAGSTM GAASITLTVQ ARQLLSGIVR QQSNLLKAIE
A_SE_SE725  AVGLGALFIG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
A_SE_SE753  AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
```

```
A_SE_SE853   AIGIGAVFIG .FLGAAGSTM GAASITLTVQ ARQLLSXIVQ QQSNLLRAIE
A_SE_SE889   AIGIGAVFIG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLMAIE
A_SE_UGSE8   AVGLAAVFFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
A_UG_92UG0   AVTLGAVFIG .FLGTAGSTM GAASITLTVQ ARKLLSGIVQ QQSNLLRAIE
A_UG_U455_   AVGLGAIFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
AC_IN_2130   AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
AC_RW_92RW   AVGLGAVFIG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
AC_SE_SE94   AVGMGAVFIG .FLGAAGSTM GAASVTLTVQ ARQLLSGIVR QQSNLLRAIE
ACD_SE_SE8   AVGIGAVFLG .FLGAAGSAM GAAAATLTVQ ARQLLSGIVQ QQSNLLKAIE
ACG_BE_VI1   AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
AD_SE_SE69   AG.LGAVFIG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLKAIE
AD_SE_SE71   AVGIGVVFFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
ADHK_NO_97   AVGMGAFFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLKAIQ
ADK_CD_MAL   IG.LGAMFLG .FLGAAGSTM GAASLTLTVQ ARQLLSGIVQ QQNNLLRAIE
AG_BE_VI11   VGLG.AVFLG .FLGAAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLRAIE
AG_NG_92NG   VGLG.AVFLG .FLGAAGSTM GAGSITLTVQ VRQLLSGIVQ QQSNLLRAIE
AGHU_GA_VI   IVGVGAVFLG .FLGVAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLKAIE
AGU_CD_Z32   AIGMGAFFLG .FLGAAGSTM GAASITLTVH VRQLLSGIVQ QQSNLLRAIE
AJ_BW_BW21   VGIMGAMFLG .FLGTAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLKAIE
B_AU_VH_AF   VGMIGAMILG .FLGAAGSTM GAASLALTVQ TRQLLSGIVQ QQNNLLRAIE
B_CN_RL42_   VGTIGAMFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQRNLLRAIE
B_DE_D31_U   VGLLGAVFLG .FLGAAGSTM GARSMALTVQ ARQLLSGIVQ QQNNLLRAIE
B_DE_HAN_U   VGMLGAMFLG .FLGAAGSTM GARSLTLTVQ ARQLLSGIVQ QQNNLLRAIE
B_FR_HXB2_   VG.IGALFLG .FLGAAGSTM GAASMTLTVQ ARQLLSGIVQ QQNNLLRAIE
B_GA_OYI__   VGMLGAMFLG .FLGAAGSTM GARSMTLTVQ ARQLLSGIVQ QQNNLLRAIE
B_GB_CAM1_   VGAIGALFLG .FLGAAGSTM GAVALTLTVQ TRQLLSGIVQ QQNNLLRAIE
B_GB_GB8_C   VGMIGAMFLG .FLGAAGSTM GAASLTLTVQ ARLLLSGIVQ QQNNLLRAIE
B_GB_MANC_   VGMLGAMFLG .FLGAAGSTM GARSITLTVQ ARQLLSGIVQ QQNNLLRAIE
B_KR_WK_AF   VT.FGALFLG .FLGAAGSTM GCTSMTLTVQ ARLLLSGIVQ QQNNLLRAIE
B_NL_3202A   VG.IGALFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
B_TW_TWCYS   VG.IGALFLG .FLGAAGSTM GAASLTLTVQ ARQLLSGIVQ QQNNLLRAIE
B_US_BC_L0   VG.IGALFLG .FLGAAGSTM GAASMTLTVQ ARQLLSGIVQ QQNNLLRAIE
B_US_DH123   VG.IGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
B_US_JRCSF   AVGIGALFLG .FLGAAGSTM GARSMTLTVQ ARQLLSGIVQ QQNNLLRAIE
B_US_MNCG_   RAAIGALFLG .FLGAAGSTM GAASVTLTVQ ARLLLSGIVQ QQNNLLRAIE
B_US_P896_   AVGIGAVFLG .FLGAAGSTM GAASVTLTVQ ARLLLSGIVQ QQNNLLRAIE
B_US_RF_M1   VGTIGAMFLG .FLGAAGSTM GAGSITLTVQ ARHLLSGIVQ QQNNLLRAIE
B_US_SF2_K   VGIVGAMFLG .FLGAAGSTM GAVSLTLTVQ ARQLLSGIVQ QQNNLLRAIE
B_US_WEAU1   VGMLGAMFLG .FLGAAGSTM GAASMTLTVQ ARLLLSGIVQ QQNNLLRAIE
B_US_WR27_   VGVIGVMFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
B_US_YU2_M   VG.LGALFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
BF1_BR_93B   VG.MGALFLG .FLGAAGSTM GAASITLTAQ ARQLLSGIVQ QQNNLLRAIE
C_BR_92BR0   AVGIGAVFLG .FLGAAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLRAIE
C_BW_96BW0   AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
C_BW_96BW1   AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARRLLSGIVQ QQSNLLRAIE
C_BW_96BW1   AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
C_BW_96BW1   AVGIGAVIFG .FLGAAGSTM GAASITLTAQ ARQLLSGIVQ QQSNLLRAIE
C_ET_ETH22   RAALGALFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLKAIE
C_IN_93IN1   AVGIGAVFLG .FLGAAGSTM GAASITLTAQ ARQLLSGIVQ QQSNLLRAIE
C_IN_93IN9   AVGIGALFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
C_IN_93IN9   VVGLGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
C_IN_94IN1   AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
C_IN_95IN2   AVGIGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF01_AE_C   AVGIGAMIFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF01_AE_C   AVGIGAMIFG .FLGAAGSTM GAASLTLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF01_AE_C   AVGIGAMIFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF01_AE_T   AVGIGAMIFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRTIE
CRF01_AE_T   AVGIGAMIFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF01_AE_T   AVGIGAMIFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF01_AE_T   AVGIGAMIFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF01_AE_T   AVGIGALIFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF01_AE_T   AVGIGAMIFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
```

```
CRF01_AE_T  AVGIGAMIFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF02_AG_F  AVGLGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF02_AG_F  AVGLGAVFFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF02_AG_G  AVGLGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF02_AG_N  AVGLGAVFLG .FLGAAGSTM GARSITLTVQ ARQLLSGIVQ QQSNLLKAIE
CRF02_AG_S  LVGLGAFFFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLKAIE
CRF02_AG_S  AVGLGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLKAIE
CRF03_AB_R  VG.IGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
CRF03_AB_R  VG.IGAVFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
CRF04_cpx_  AVGIGAMFLG .FLGAAGSTM GAASMTLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF04_cpx_  AVGLGALFLG .FLGAAGSTM GAASLTLTVQ ARQLLSGIVQ QQSNLLRAIE
CRF04_cpx_  AVGIGAVFPG LFLGAAGSTM GAASITLTVQ ARQLLFGIVQ QQSNLLRAIE
CRF05_DF_B  MG.IGAMFLG .FLAAAGSTM GAASIALTVQ ARQLLSGIVQ QQNNLLQAIE
CRF05_DF_B  VG.VGALLIG .FLGAAGSTM GAASMTLTVQ ARQLLSGIVQ QQNNLLQAIE
CRF06_cpx_  VGLG.AMFLG .FLGTAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLRAIE
CRF06_cpx_  VGLG.AVFLG .FLGTAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLRAIE
CRF06_cpx_  VGIG.AFFLG .FLGTAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLRAIE
CRF06_cpx_  VGLG.AVFLG .FLGTAGSTM GAASITLTVQ VRQLLSGIVH QQSNLLRAIE
CRF11_cpx_  VGIG.AVLFG .FLGAAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLKAIE
CRF11_cpx_  VGIG.AVLLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLKAIE
D_CD_84ZR0  IG.LGAMFLG .FLGAAGSTM GAASMTLTVQ ARQVLSGIVQ QQNNLLRAIE
D_CD_ELI_K  IG.LGAMFLG .FLGAAGSTM GARSVTLTVQ ARQLMSGIVQ QQNNLLRAIE
D_CD_NDK_M  IG.LGAVFLG .FLGAAGSTM GAASVTLTVQ ARQLMSGIVH QQNNLLRAIE
D_UG_94UG1  IG.LGALFLG .FLGTAGSTM GAVSLTLTVQ ARQVLSGIVQ QQNNLLRAIE
F1_BE_VI85  AG.LGALFLG .FLGDSREHM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
F1_BR_93BR  VG.LGALFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
F1_FI_FIN9  VA.IGAVFLG .FLSAAGSTM GAASLTLTVQ ARQLLSGIVQ QQNNLLQAIE
F1_FR_MP41  VG.IGALFLR .FLGAAGSNI GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
F2_CM_MP25  VG.MGAMFLG .FLGAAGSTM GAASITLTVQ ARNLLSGIVQ QQSNLLKAIE
F2KU_BE_VI  AG.LGALFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSDLLRAIE
G_BE_DRCBL  VGVG.AIFLG .FLGTAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLRAIE
G_NG_92NG0  VGLG.AVFLG .FLGAAGSTM GAASITLTAQ VRQLLSGIVQ QQSNLLRAIE
G_SE_SE616  VGLG.AVFLG .FLGAAGSTM GAASITLTVQ VRQLLSGIVQ QQGNLLRAIE
H_BE_VI991  AVGMGAFFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIQ
H_BE_VI997  AVGMGAFFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIQ
H_CF_90CF0  AVGMGASFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIQ
J_SE_SE702  VGIG.AVFLG .FLGTAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLKAIX
J_SE_SE788  VGIG.AVFLG .FLGTAGSTM GAASITLTVQ VRQLLSGIVQ QQSNLLKAIE
K_CD_EQTB1  VG.IGALFLG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQNNLLRAIE
K_CM_MP535  VG.LGAVFFG .FLGAAGSTM GAASITLTVQ ARQLLSGIVQ QQSNLLRAIE
N_CM_YBF30  AFGLGALFLG .FLGAAGSTM GAASITLTVQ ARTLLSGIVQ QQNILLRAIE
O_CM_ANT70  AVGLGMLFLG .VLSAAGSTM GAAATTLAVQ THTLLKGIVQ QQDNLLRAIQ
O_CM_MVP51  AVGLGMLFLG .VLSAAGSTM GAAATALTVR THSVLKGIVQ QQDNLLRAIQ
O_SN_99SE_  AVGLGMLFLG .VLSAAGSTM GAAATALAVQ TQSLMKGIVQ QQDNLLRAIQ
O_SN_99SE_  AVGLGMLFLG .VLSAAGSTM GAAATTLAVQ THTLMKGIVQ QQDNLLRAIQ
U_CD___83C  AVGMGALFLG .FLGAAGSTM GAASMALTAQ ARQLLSGIVQ QQNNLLRAIE
```

```
           651                                                        700
00BW0762_1 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW0768_2 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGLWGCSG KLICTTSVHW
00BW0874_2 AQQHMLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTAVPR
00BW1471_2 AQQHMLQLTV WGIKQLQARV LALERYLQDQ QLLGIWGCSG KLICTTPVPW
00BW1616_2 AQQHMLQLTV WGIKQLQARV LAIERYLRDQ QLLGIWGCSG KLICTTNVPW
00BW1686_8 AQQHMLQLTV WGIKQLQTRV LAMERYLKDQ QLLGLWGCSG KIICTTNVPW
00BW1759_3 AQQHMLQLTV WGINQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW1773_2 AQQHLLQLTV WGIKQLQTRV LSIERYLQDQ QLLGIWGCSG KLICTTAVPW
00BW1783_5 AQQHMLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW1795_6 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW1811_3 AQQHLLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW1859_5 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW1880_2 AQQHMLQLTV WGIKQLQARV LAIEKYLKDQ QLLGMWGCSG KIICTTTVPW
00BW1921_1 AQQHLLQLTV WGIKQLQTRV LAIERYLKDQ QILGIWGCSG KLICTTSVPW
00BW2036_1 AQQHMLQLTV WGIKQLQARV LAIERYLRDQ QLLGLWGCSG KLICTTTVPW
00BW2063_6 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICPTAVPW
00BW2087_2 AQQHLLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW2127_2 AQQHLLQLTV WGIKQLQTRV LAIESYLKDQ RLLGIWGCSG KLICTTAVPW
00BW2128_3 ARRHLLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW2276_7 AQQHMLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW3819_3 AQQHLLQLTV WGIKQLQTRV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
00BW3842_8 AQQHMLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTNVPW
00BW3871_3 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGLWGCSG KLICTTNVPW
00BW3876_9 AQQHLLQLTV WGIKQLQTRV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
00BW3886_8 AQQHLLQLTV WGIKQLQARV LAMERYLQDQ QLLGIWGCSG KLICTTAVPW
00BW3891_6 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTNVPW
00BW3970_2 AQHHLLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
00BW5031_1 AQQHMLQLTV WGIKQLQARV LAIERYLRDQ QLLGIWGCSG KIICTTAVPW
 96BW01B21 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTNVPW
  96BW0407 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTNVPW
  96BW0502 AQQHLLQLTV WGIKQLQTRI LAVERYLKDQ QLLGIWGCSG KLICTTAVPW
 96BW06_J4 AQQHMLQLTV WGIKQLQTKV LAIERYLKDQ QLLGFWGCSG KLVCTTAVPW
 96BW11_06 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
  96BW1210 AQQHMLQLTV WGIKQLQARV LALERYLRDQ QLLGIWGCSG KLICTTNVPW
 96BW15B03 AQQHMLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTTVPW
 96BW16_26 AQQHLLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
 96BW17A09 AQQHMLQLTV WGIKQLQARV LALERYLKGQ QLLGIWGCSG KLICTTSVPW
 96BWMO1_5 AQQHMLQLTI WGIKQLQTRV LAVERYLRDQ QLLGIWGCSG KLICTTAVPW
 96BWMO3_2 AQQHMLQLTV WGIKQLRARV LAIERYLKDQ QLLGVWGCSG RLICTTAVPW
98BWMC12_2 AQQHMLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTNVPW
98BWMC13_4 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
98BWMC14_a AQQHLLQLTV WGIKQLQTRV LAIERYLKDQ QLLSLWGCSG RLICTTNVPW
98BWMO14_1 AQQHMLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
98BWMO18_d AQQHMLQLTV WGIKQLQARV LALERYLRDQ QLLGIWGCSG KLICTTNVPW
98BWMO36_a AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
98BWMO37_d AQQHLLQLTV WGIKQLQARV LAMERYLKDQ QLLGIWGCSG KLICTTAVPW
99BW3932_1 AQQHLLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
99BW4642_4 AQQHMLQLTV WGIKQLQARV LAIERYLKDQ QLLGLWGCSG KLICTTAVPW
99BW4745_8 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KIICTTAVPW
99BW4754_7 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
99BWMC16_8 AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGYSG KLICTTTVPW
A2_CD_97CD AQQQMLRLTV WGIKQLQARV LALERYLQDQ QLLGIWGCSG KLICATDVRW
A2_CY_94CY AQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICATTVPW
A2D___97KR AQQHMLRLTV WGIKQLQARV LAVERYLQDQ QLLGIWGCSG KLICTTFVPW
A2G_CD_97C AQQHLLKLTV WGIKQLQARV LALERYLQDQ QLLGLWGCSG KLICTTIVPW
A_BY_97BL0 AQQXLLKLTV XGIKQLQARX LAVEXYLKDQ QXLRIXGCSX KLICTTNVPX
A_KE_Q23_A AQQHLLKLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTNVPW
A_SE_SE659 AQQHLLKLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
A_SE_SE725 AQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
A_SE_SE753 VQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTTVPW
```

```
A_SE_SE853   AQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
A_SE_SE889   AQQHLLKLTV WGIKQLQAQV LAVERYLRDQ QLLGIWGCSG KLICTTTVPW
A_SE_UGSE8   AQQHMLRLTV WGIKQLQARV MAVERYLKDQ QLLGIWGCSG KIICTTAVPW
A_UG_92UG0   AQQHLLKLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICPTNVPW
A_UG_U455_   AQQHLLKLTV WGIKQLQARV LAVERYLQDQ QLLGIWGCSG KLICTTTVPW
AC_IN_2130   AQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
AC_RW_92RW   AQQHLLKLTV WGIKQLQARV LALERYLRDQ QLLGIWGCSG KLICTTNVPW
AC_SE_SE94   AQQHLLKLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KIICTTNVPW
ACD_SE_SE8   AQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICATNVPW
ACG_BE_VI1   AQQHMLQLTV WGIKQLQTRV LAIERYLQVQ QLLGIWGCSG KLICTTSVPW
AD_SE_SE69   AQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
AD_SE_SE71   AQQHLLKLTV WGIKQLQARV LALERYLKDQ QLLGIWGCSG KLICPTTVPW
ADHK_NO_97   AQQHMLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
ADK_CD_MAL   AQQHLLQLTV WGIKQLQARV LAVERYLQDQ RLLGMWGCSG KHICTTFVPW
AG_BE_VI11   AQQHMLQLTV WGIKQLQARV LAVERFLKDQ QLLGIWGCSG KLICTTNVPW
AG_NG_92NG   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
AGHU_GA_VI   AQQHLLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTNVPW
AGU_CD_Z32   AQQHLLKLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
AJ_BW_BW21   AQQHLLRLTV WGIKQLQAKI LAVERYLQDQ QLLGIWGCSG KIICPTNVPW
B_AU_VH_AF   AQQHLLQLTV WGIKQLQARI LAVERYLEDQ QLLGIWGCSG KFICTTTVP.
B_CN_RL42_   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTSVPW
B_DE_D31_U   AQQHLLQLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTAVPW
B_DE_HAN_U   AQQHLLQLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTAVPW
B_FR_HXB2_   AQQHLLQLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KLICTTTVPW
B_GA_OYI__   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTAVPW
B_GB_CAM1_   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTTVPW
B_GB_GB8_C   AQQHLLRLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTAVPW
B_GB_MANC_   AQQHLLQLTV WGIKQLQARV LAVERYLQDQ QLLGIWGCSG KLICTTTVPW
B_KR_WK_AF   AQQHLLQLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTAVPW
B_NL_3202A   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
B_TW_TWCYS   AQQHMLQLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTAVPW
B_US_BC_L0   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICNTNVPW
B_US_DH123   AQQHMLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTAVPW
B_US_JRCSF   AQQHMLQLTV WGIKQLQARV LAVERYLQDQ QLLGIWGCSG KLICTTTVPW
B_US_MNCG_   AQQHMLQLTV WGIKQLQARV LAVERYLKDQ QLMGIWGCSG KLICTTAVPW
B_US_P896_   AQQHMLQLTV WGIKQLQARV LAVERYLKDQ QLLGFWGCSG KLICTTTVPW
B_US_RF_M1   AQQHLLQLTV WGIKQLQARV LALERYLRDQ QLMGIWGCSG KLICTTSVPW
B_US_SF2_K   AQQHLLQLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTTVPW
B_US_WEAU1   AQQHLFELTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTAVPW
B_US_WR27_   AQQHMLQLTA WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTTVPW
B_US_YU2_M   AQQHLLQLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTAVPW
BF1_BR_93B   AQQHLLQLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTTVPW
C_BR_92BR0   AQQHMLQLTV WGIKQLQARI LAVERYLKDQ QLLGLWGCSG KLICTTDVPW
C_BW_96BW0   AQQHMLQLTV WGIKQLQTRV LAIERYLRDQ QLLGIWGCSG KLICTTAVPW
C_BW_96BW1   AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTNVPW
C_BW_96BW1   AQQHMLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
C_BW_96BW1   AQQHMLQLTV WGIKQLQARV LALERYLRDQ QLLGIWGCSG KLICTTNVPW
C_BW_96BW1   AQQHMLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTTVPW
C_ET_ETH22   AQQHMLQLTV WGIKQLQTRV LAIERHLRDQ QLLGIWGCSG KLICTTAVPW
C_IN_93IN1   AQQHLLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
C_IN_93IN9   AQQHLLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
C_IN_93IN9   AQQHLLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
C_IN_94IN1   AQQHLLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
C_IN_95IN2   AQQHLLQLTV WGIKQLQTRV LAIERYLKDQ QLLGIWGCSG KLICTTAVPW
CRF01_AE_C   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ KFLGLWGCSG KIICTTNVPW
CRF01_AE_C   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ KFLGLWGCSG KIICTTSVPW
CRF01_AE_C   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QFLGLWGCSG KIICTTSVPW
CRF01_AE_T   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ KFLGLWGCSG KIICTTAVPW
CRF01_AE_T   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ KFLGLWGCSG KIICTTAVPW
CRF01_AE_T   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ KFLGLWGCSG KIICTTAVPN
CRF01_AE_T   AQQHMLQLTV WGIKQLQARV LAVERYLKDQ KFLGLWGCSG KIICTTAVPW
CRF01_AE_T   AQQHLLQLTV WGIKQLQARV LAVERYLKDQ KFLGLWGCSG KIICTTAVPW
```

```
CRF01_AE_T  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ KFLGLWGCSG KIVCTTAVPW
CRF02_AG_F  AQQHLLKLTV WGIKQLQARV LALESYLRDQ QLLGIWGCSG KLICTTNVPW
CRF02_AG_F  AQQHLLKLTV WGIKQLQARV LALERYLRDQ QLLGIWGCSG KLICTTTVPW
CRF02_AG_G  AQQHLLKLTV WGIKQLQARV LALERYLSDQ QLLGIWGCSG KLICTTNVPW
CRF02_AG_N  AQQHLLKLTV WGIKQLQARV LALERYLRDQ QLLGIWGCSG KLICTTTVPW
CRF02_AG_S  AQQHLLRLTV WGIKQLQARV LALEAYLKDQ QLLGIWGCSG KLICTTTVPW
CRF02_AG_S  AQQHLLKLTV WGIKQLQARV LALERYLKDQ QLLGIWGCSG KLICTTTVPW
CRF03_AB_R  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTAVPW
CRF03_AB_R  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTAVPW
CRF04_cpx_  AQQHLLRLTV WGIKQLQARV LALESYLKDQ QLLGIWGCSG KLICTTNVPW
CRF04_cpx_  AQQHLLKLTV WGIKQLQARV LALESYLKDQ QLLGIWGCFG KLICTTNVPW
CRF04_cpx_  AQQQLLRLTV WGVKQLQARV LALESYLRDQ QLLGIWGCSG RLICTTNVPW
CRF05_DF_B  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGLWGSSG KHICSTSVPW
CRF06_cpx_  AQQHLLQLTV WGIKQLQARI LAVERYLRDQ QLLGLWGCSG KLICTTSVPW
CRF06_cpx_  AQQHLLQLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KLICPTNVLW
CRF06_cpx_  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICPTNVPW
CRF06_cpx_  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICPTNVPW
CRF11_cpx_  AQQHLLKLTV WGIKQLQARV LALERYLKDQ QLLGIWGCSG KLICPTNVPW
CRF11_cpx_  VQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
D_CD_84ZR0  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
D_CD_ELI_K  AQQHLLQLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KHICTTTVPW
D_CD_NDK_M  AQQHLLQLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KHICTTNVPW
D_UG_94UG1  AQQHLLQLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG RHICTTNVPW
F1_BE_VI85  AQQHLLQLTV WGIKQLQARI LAVESYLKDQ QLLGIWGCSG KHICTTNVPW
F1_BR_93BR  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
F1_FI_FIN9  AQQHMLQLTV WGIKQLQARV LAVERYLKDQ QLLGLWGCSG KLICTTNVPW
F1_FR_MP41  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGLWGCSG KLICTTNVPW
F2_CM_MP25  AQQHLLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
F2KU_BE_VI  AQQHLLQLTV WGIKQLQARI LAVERYLKDQ QLLGIWGCSG KLICPTTVPW
G_BE_DRCBL  AQQHLLQLTV WGIKQLQARV LAIERYLKDQ QLLGIWGCSG KLICTTNVPW
G_NG_92NG0  AQQHLLQLTV WGIKQLRARV LALERYLKDQ QLLGIWGCSG KLICTTNVPW
G_SE_SE616  AQQHLLQLTV WGIKQLQSRV LAIERYLKDQ QLLGIWGCSG KLICTTNVPW
H_BE_VI991  AQQHLLQLTV WGIKQLQARL LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
H_BE_VI997  AQQHMLQLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
H_CF_90CF0  AQQHMLQLTV WGVKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
H_CF_90CF0  ARQHMLQLTV WGIKQLQARV LAVERYLRDQ QLLGIWGCSG KLICTTNVPW
J_SE_SE702  AQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
J_SE_SE788  AQQHLLKLTV WGIKQLQARV LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
K_CD_EQTB1  AQQQMLQLTV WGIKQLRARV LAVERYLRDQ QLLGIWGCSG KLICTTNVPW
K_CM_MP535  AQQHLLQLTV WGIKQLRARI LAVERYLKDQ QLLGIWGCSG KLICTTNVPW
N_CM_YBF30  AQQHLLQLSI WGIKQLQAKV LAIERYLRDQ QILSLWGCSG KTICYTTVPW
O_CM_ANT70  AQQQLLRLSX WGIRQLRARL LALETLLQNQ QLLSLWGCKG KLVCYTSVKW
O_CM_MVP51  AQQHLLRLSV WGIRQLRARL QALETLIQNQ QRLNLWGCKG KLICYTSVKW
O_SN_99SE_  AQQQLLRLSV WGIRQLRARL LALETLIQNQ QLLNLWGCKG KLICYTSVKW
O_SN_99SE_  AQQQLLRLSV WGIRQLRARL LALETLIQNQ QLLNLWGCKG RLVCYTSVKW
U_CD___83C  AQQHLLQLTV WGIKQLQARV LAVERYLESQ QLLGLWGCSG KLICTTTVPW
```

```
            701                                                   750
00BW0762_1  NASWSN.... KSEGDIWNN. .MTWMQWDRE ISNYTNTIYR LLENSQIQQE
00BW0768_2  NSSWSN.... KSQKEIWDN. .MTWMQWDKE ISNYTDTIYR LLEESQSQQE
00BW0874_2  NNSWSN.... KTLGDIWEN. .MTWMQWDRE ISNYTDTIYR LLEVSQNQQE
00BW1471_2  NSSWSN.... KSVDDIWKN. .MTWMEWDKE INNYTKIIYN LLEKSQKQQE
00BW1616_2  NSSWST.... RNYSDIWNN. .MTWMQWDRE IDNYTDIIYR LLEESQNQQE
00BW1686_8  NASWSN.... KSQEDIWNN. .MTWMPWDRE INNYTRTIYE LLEESQSQQE
00BW1759_3  NHSWSN.... KSEEDIWNH. .TTWMQWDRE ISNYTDTIYR LLEDSQNQQE
00BW1773_2  NSSWSN.... RSQDDIWEN. .MTWMQWDKE ISNYTNTIYN LLEKSQNQQE
00BW1783_5  NTSWSN.... KSQNDIWDN. .MTWMQWDRE ISNYTGIIYN LLEESQIQQE
00BW1795_6  NVSWSN.... KSYKAIWDN. .MTWMQWDRE ISNYTDTIYS LLEDSQNQQE
00BW1811_3  NNSWSN.... KSFTEIWDN. .MTWMEWDRE ISNYTNTIYR LLEKSQNQQE
00BW1859_5  NSSWSN.... KSQADIWDN. .MTWMQWDRE ISNYTDTIYR LLEDSQIQQE
00BW1880_2  NDSWSN.... KSQEYIWGN. .MTWMQWDRE INNYTNTIYR LLENSQNQRE
00BW1921_1  NSSWSN.... KSLTEIWDN. .MTWMQWDRE ISNYTETIYR LLEDSQNQQE
00BW2036_1  NSSWSN.... KSLADIWDN. .MTWMQWDKE INNYTGIIYK LLEDSQIQQE
00BW2063_6  NASWSN.... KSHDEIWEN. .MTWMQWDRE ISNYTDIIYR MLEDSQNQQE
00BW2087_2  NSSWSN.... KSQAEIWDN. .MTWMQWDRE ISNYTNTIYK LLEDSQLQQE
00BW2127_2  NSSWSS.... RSHEDIWNN. .LTWMQWDRE IDNYTDTIYK LLEESQNQQE
00BW2128_3  NSSWSN.... RSKTEIWDN. .MTWMQWDRE ISNYTETIYR LLEESQNQQE
00BW2276_7  NSSWSN.... KSQTDIWGN. .MTWVQWDKE INNYTDTIYR LLEEWQNQQE
00BW3819_3  NSSWSN.... KTNSEIWDN. .MTWMQWDRE VINYTGIIYG LLEESQNQQE
00BW3842_8  NSSWSN.... KSLEDIWDN. .MTWMQWDRE ISNYTREIYK LLEESQNQQE
00BW3871_3  NLSWSN.... KSQKDIWES. .MTWMQWDNE ISNYTGIIYE LLEDSQNRQE
00BW3876_9  NSSWSN.... KSQSDIWEN. .MTWMQWDRK ISNYTTTIYR LLEESQTQQE
00BW3886_8  NSSWSN.... KTQEDIWNK. .TTWMQWDKE ISNYTDIIYS LLADSQNQQE
00BW3891_6  NASWSN.... KSQEDIWNN. .MTWMQWDKE VSNYTNTIYK LLEESQNQQE
00BW3970_2  NSSWSN.... KSQADIWDN. .MTWMQWDRE ISNYTDTIYR LLEVSQTQQE
00BW5031_1  NSSWSN.... KNEEDIWGN. .MT.MQWDRE ISNYTDTIYR LLEDSQSQQE
 96BW01B21  NSSWSN.... RSEADIWDS. .MTWMQWDKE ISNYTGTIYR LLEESQTQQE
  96BW0407  NNSWSN.... KSLNYIWEN. .MTWMEWDRE ISNYSDTIYR LLEDSQNQQE
  96BW0502  NSSWSN.... RSHDEIWDN. .MTWMQWDRE INNYTDTIYR LLEESQNQQE
 96BW06_J4  NSSWSN.... KSLGDIWDN. .MTWMQWDRE ISNYTGTIYR LLEDSQNQQE
 96BW11_06  NASWSN.... KSQEEIWGN. .MTWMQWDRE ISNYTDTIYR LLEVSQNQQE
  96BW1210  NSSWSN.... KTENEIWEN. .MTWMQWERE IDNYTDTIYR LLEVSQNQQE
 96BW15B03  NSSWSN.... KTQGEIWEN. .MTWMQWDKE ISNYTGIIYR LLGESQNQQE
 96BW16_26  NVSWSN.... RSEDDIWNN. .MTWMQWDRE INNYTNTIYR LLEESQNQQE
 96BW17A09  NSTWSN.... KSLDDIWQN. .MTWMEWDRE INNYTNTIYR LLEESQNQQE
 96BWMO1_5  NSSWSN.... KSQHDIWDN. .LTWMQWDRE ISNYTDTIYR LLEESQNQQE
 96BWMO3_2  NTSWSN.... KSLTDIWDN. .MTWMQWERE IDNYTGIIYS LLEESQIQQE
98BWMC12_2  NSSWSN.... KSQTDIWDK. .MTWMEWDRE INNYTDTIYW LLEKSQNQQE
98BWMC13_4  NASWSN.... KSQGEIWEN. .MTWMQWDRE INNYTDTIYR LLEESQNQQE
98BWMC14_a  NSSWSN.... KSYEDIWDN. .MTWMQWDKE ISNYTDTIYR LLEVSQNQQE
98BWMO14_1  NSSWSN.... KSYGDIWDN. .MTWMQWDRE ISNYTDTIYK LLEDSQIQQE
98BWMO18_d  NSSWSN.... RTQDDIWGN. .MTWMQWDRE INNYTDIIYG LLEESQYQQE
98BWMO36_a  NSSWSN.... KSYDSIWDN. .MTWMQWDRE INNYTDTIYK LLEESQNQQE
98BWMO37_d  NSSWSN.... KTQNDIWEN. .MTWMQWDRE ISNYTDTIYR LLEESQTQQE
99BW3932_1  NSSWSN.... KSQEEIWNN. .MTWMQWDRE VSNYTDTIYR LLEVSQNQQE
99BW4642_4  NSSWSN.... KSKEDIWNN. .MTWMQWDRE ISNYTNTIYR LLEDSQSQQE
99BW4745_8  NNSWSN.... KTEQQIWNS. .TTWMEWDRE ISNYTDIIYR LLENSQYQQE
99BW4754_7  NTSWSN.... KSEEEIWEN. .MTWMQWDRE INNYTDLIYR MLEESQNQQE
99BWMC16_8  NDSWSN.... KSKEEIWGN. .MTWMQWDKE INNYTSLIYT LLEDSQIQQE
A2_CD_97CD  NSSWSN.... KTQEQIWKN. .MTWLQWDKE ISTYTDIIYM LLEESQNQQE
A2_CY_94CY  NTSWSN.... KSQDEIWDN. .MTWLQWDKE ISNYTNIIYR LLEESQNQQE
A2D___97KR  NSSWSN.... RTQDDIWNN. .MTWLQWERE IDNYTGTIYR LLEESQNQQE
A2G_CD_97C  NSSWSN.... KTQNEIWDN. .MTWLQWDKE INNYTDTIYK LLEDSQNQQE
A_BY_97BL0  XSSXSN.... KXQSEIXDN. .MTMQDX..E VINXTDIIYD LIXKSQNQQE
A_KE_Q23_A  NSSWSN.... KSLDEIWNN. .MTWLQWDKE INNYTQLIYR LIEESQNQQE
A_SE_SE659  NTSWSN.... KTQSEIWGN. .MTWLQWDKE ISKYTGIIYD LLEESQNQQE
A_SE_SE725  NSSWSN.... KSQNEIWED. .MTWLQWDRE ISNYTEIIYK LIEESQNQQE
A_SE_SE753  NSSWSN.... KSYDEIWDN. .MTWLQWDKE ISNYTEIIYQ LIEKSQNQQE
```

```
A_SE_SE853   NSSWSN.... KSQSEIWDN. .MTWLQWDKE ISNYTQTIYR LIEESQNQQE
A_SE_SE889   NSSWSD.... VSQSEIWEN. .MTWLQWDKE ISNYTQIIYS LIEESQNQQE
A_SE_UGSE8   NSTWSN.... KSYTQIWDN. .MTWLQWDKE ISNYTDIIYQ LIEESQYQQE
A_UG_92UG0   NSSWSN.... KSLDEIWEN. .MTWLQWDKE ISNYTIKIYE LIEESQIQQE
A_UG_U455_   NSSWSN.... KSQEDIWNN. .MTWLQWEKE ISSYTGIIYQ LIEESQNQQE
AC_IN_2130   NASWSN.... KSQKEIWEN. .MTWLQWDKE ISNYTDIIYH LIKESQNQQE
AC_RW_92RW   NSSWSN.... KTQQEIWDN. .MTWQQWDKE IGNYTQIIYS LIEESQNQQE
AC_SE_SE94   NSSWSN.... KSLSKIWDN. .MTWLEWDKE ISNYTQIIFD LIEESQIQQE
ACD_SE_SE8   NSSWSN.... KSQAEIWQN. .MTWLQWDKE IDNYTQIIYM LLEEPQNQQE
ACG_BE_VI1   NSSWSN.... KSQGEIWDN. .MTWMQWDRE ISNYSDTIYR LLEESQNQQE
AD_SE_SE69   NSTWSN.... KSQEEIWDN. .MTWLQWDKE ISNYTQTIYN LLAESQNQQE
AD_SE_SE71   NSSWSN.... KSWSEIWNN. .MTWLQWDKE INNYTHIIYQ LIEESQNQQE
ADHK_NO_97   NSSWSN.... KSLNDIWXN. .MTWMEWDKQ ISNYSEEIYR LLEVSQIQQE
ADK_CD_MAL   NSSWSN.... RSLDDIWNN. .MTWMQWEKE ISNYTGIIYN LIEESQIQQE
AG_BE_VI11   NVSWSN.... KSYNDIWNN. .MTWIEWERE ISNYTQQIYN LLEESQNQQE
AG_NG_92NG   NTSWSN.... KSYEEIWDN. .MTWIQWERE VSNYTQQIYS LIEESQNQQE
AGHU_GA_VI   NSSWSN.... KSQEDIWSN. .MTWMQWEKE ISNYTNTIYW LLEESQNQQE
AGU_CD_Z32   NSSWSN.... KSQSDIWDK. .MTWLEWDKE VSNYTQVIYN LLEESQTQQE
AJ_BW_BW21   NSSWSN.... KSYNEIWEN. .MTWRDWERE IDKYTDTIYS LIGEAQSQQD
B_AU_VH_AF   NASWSN.... KSLSDIWDN. .MTWMQWERE IGNYTGLIYH LLEESQNQQE
B_CN_RL42_   NASWSN.... KSLHEIWNN. .MTWMEWERE IDNYTREIYT LIEESQNQQE
B_DE_D31_U   NASWSN.... KSMDMIWNN. .MTWMEWERE IDNYTSLIYT LIEESQNQQE
B_DE_HAN_U   NASWSN.... KTLDQIWNN. .MTWMEWERE IDNYTSLIYT LIEQSQNQQE
B_FR_HXB2_   NASWSN.... KSLEQIWNH. .TTWMEWDRE INNYTSLIHS LIEESQNQQE
B_GA_OYI__   NASWSN.... KSLNEIWDN. .MTWMQWERE IDNYTHLIYT LIEESQNQQE
B_GB_CAM1_   NASWSN.... KSLDKIWNN. .MTWMEWERE IDNYTNLIYT LIEESQNQQE
B_GB_GB8_C   NASWSS.... KSLDKIWNN. .TTWMEWERE IDNYTNTTYT LIEESQNQQE
B_GB_MANC_   NSSWSN.... KSLEQIWGN. .MTWMQWEKE INNYTGLIYT LIQESQNQQE
B_KR_WK_AF   NTSWSN.... KSENEIWDN. .MTWMEWDRE INNYTNLIYD LLEKSQNQQE
B_NL_3202A   NASWSN.... KSLDKIWNN. .MTWMEWDRE VSNYTSLIYT LIEESQNQQE
B_TW_TWCYS   NISWSN.... KTYDEIWNN. .MTWMQWERE IDNYTDHIYN LLEKSQNQQE
B_US_BC_L0   NASWSN.... KSLNQIWDN. .MTWMQWERE IENYTDLIYN LIEESQNQQE
B_US_DH123   NTSWSN.... KSLDTIWGN. .MTWMQWEKE INNYTGLIYN LIEESQNQQE
B_US_JRCSF   NTSWSN.... KSLDSIWNN. .MTWMEWEKE IENYTNTIYT LIEESQIQQE
B_US_MNCG_   NASWSN.... KSLDDIWNN. .MTWMQWERE IDNYTSLIYS LLEKSQTQQE
B_US_P896_   NVSWSN.... KSVDDIWNN. .MTWMEWERE IDNYTDYIYD LLEKSQTQQE
B_US_RF_M1   NASWSN.... KSLNMIWNN. .MTWMQWERE IDNYTGIIYN LLEESQNQQE
B_US_SF2_K   NASWSN.... KSLEDIWDN. .MTWMQWERE IDNYTNTIYT LLEESQNQQE
B_US_WEAU1   NASWSN.... RSQDYIWNN. .MTWMEWERE INNYTGLIYN LIEESQNQQE
B_US_WR27_   NATWSN.... KSLDKIWQN. .MTWMEWERE IDNYTELIYS LLEQSQNQQE
B_US_YU2_M   NTSWSN.... KSLNEIWDN. .MTWMKWERE IDNYTHIIYS LIEQSQNQQE
BF1_BR_93B   NSSWSN.... KSQEKIWGN. .MTWMEWEKE ISNYSNEIYR LIEESQNQQE
C_BR_92BR0   NSSWSN.... RSQEDIWNN. .MTWMQWDRE ISNYTNTIYR LLEDSQNQQE
C_BW_96BW0   NNSWSN.... KSLDYIWGN. .MTWMEWDRE ISNYSNTIYR LLEDSQNQQE
C_BW_96BW1   NASWSN.... KSQEEIWGN. .MTWMQWDRE ISNYTDTIYR LLEVSQNQQE
C_BW_96BW1   NSSWSN.... KTENEIWEN. .MTWMQWERE IDNYTDTIYR LLEVSQNQQE
C_BW_96BW1   NSSWSN.... KTQGEIWEN. .MTWMQWDKE ISNYTGIIYR LLGESQNQQE
C_ET_ETH22   NSSWSN.... KSQEEIWDN. .MTWMQWDRE ISNYTDIIYN LLEVSQNQQD
C_IN_93IN1   NSSWSN.... KTQSEIWNN. .MTWMQWDRE VSNYTNIIYS LLEESQNQQE
C_IN_93IN9   NSSWSN.... RSQQDIWDN. .MTWMQWDRE ISNYTNTIYR LLEDSQNQQE
C_IN_93IN9   NSSWSN.... KSQIEIWEN. .MTWMQWDRE INNYTQTIYR LLEDSQNQQE
C_IN_94IN1   NSSWSN.... RTQEEIWNN. .MTWMQWDRE INNYTNTIYR LLEDSQNQQE
C_IN_95IN2   NSSWSN.... RTQKEIWDN. .MTWMQWDRE INNYTNTIYR LLEESQNQQE
CRF01_AE_C   NTSWSN.... KSYAEIWDN. .MTWIEWDKE INNYTNQIYE LLTKSQNQQE
CRF01_AE_C   NSSWSN.... KSYEAIWNN. .MTWIEWDRE INEYTNQIYE LLTESQDQQE
CRF01_AE_C   NSSWSN.... KSYNEIWNN. .MTWIEWDRE INNYTNQIYE LLTKSQDQQE
CRF01_AE_T   NSTWSN.... RSFEEIWNN. .MTWIEWERE ISNYTNQIYE ILTESQNQQD
CRF01_AE_T   NSTWSN.... RSYEEIWNN. .MTWTEWERE ISNYTNQIYE ILTESQDQQD
CRF01_AE_T   .STWSN.... RSYEEIWDN. .MTWIEWERE ISNYTNQIYE ILTESQNQQE
CRF01_AE_T   NSTWSN.... KSFDEIWNN. .MTWIEWERE ISNYTNKIYD ILTESQNQQD
CRF01_AE_T   NSTWSN.... RSFEEIWNN. .MTWIEWERE ISNYTNQIYE ILTESQNQQD
```

```
CRF01_AE_T  NSTWSN.... KSYEEIWDK. .MTWTQWERE ISNYTSTIYE ILTESQNQQD
CRF02_AG_F  NSSWSN.... KTYNDIWDN. .MTWLQWDKE VSNYTDIIYN LIEESQNQQE
CRF02_AG_F  NASWSN.... KTYNDIWDN. .MTWLQWDKE ISNYTDIIYN LIEESQNQQE
CRF02_AG_G  NSSWSN.... RTFEDIWEN. .MTWLQWEKE ISNYTDTIYA LIEDSQNQQE
CRF02_AG_N  NSSWSN.... KTFNDIWDN. .MTWIQWEKE ISNYTDIIYN LIEESQNRQE
CRF02_AG_S  NSSWSN.... KTYDHIWGN. .MTWLQWDKE ISNYTHIIYD LIEESQNQQE
CRF02_AG_S  NSSWSNN... KTFKDIWDN. .MTWLQWDKE ISNYTEKIYT LIEEAQNQQE
CRF03_AB_R  NTSWSN.... KPLDEINN.. .MTWMEWERE INNYTGLIYN LIEESQNQQE
CRF03_AB_R  NTSWSN.... KSLDKIWNN. .MTWMEWERE INNYTGLIYN LIEESQNQQE
CRF04_cpx_  NSSWSN.... KSYNDIWDN. .MTWLQWDKE INNYTQIIYG LLEESQNQQE
CRF04_cpx_  NSSWNN.... KSYNDIWGN. .MTWLQWDKE INNYTQIIYE LLEESQIQQE
CRF04_cpx_  NSSWSN.... KSYEKIWDN. .MTWLQWDKE INNYTEIIYD LLGEAQNQQE
CRF05_DF_B  NSSWSN.... KSEGEIWDN. .MTWMEWEKE ISNYSSTIYR LIEQSQIQQE
CRF05_DF_B  NSSWSN.... KSQEEIWEN. .MTWMQWEKE IGQYSNTIYR LIEQSQNQQE
CRF06_cpx_  NTSWSN.... KTYDEIWGN. .MTWIEWDRE INNYTQQIYS LIELSQTQQE
CRF06_cpx_  NASWSN.... KTYNEIWDN. .MTWIEWDRE INNYTQQIYS LIEESQNQQE
CRF06_cpx_  NASWSN.... RTFNEIWNN. .MTWIEWDRE INNYTQQIYS LIEESQSQQE
CRF06_cpx_  NISWSN.... KTYSEIWDN. .MTWIEWDRE INNYTQQIYN LIEESQNQQE
CRF11_cpx_  NISWSN.... KSYDEIWDN. .MTWIEWERE IDNYTQTIYT LLEESQTQQE
CRF11_cpx_  NVSWSN.... KTYNEIWENE NMTWIKWERE IDNYTQTIYT LIEESQNQQE
D_CD_84ZR0  NSSWSN.... RSVEYIWGN. .MTWMQWERE IDNYTGLIYN LIEESQIQQE
D_CD_ELI_K  NSSWSN.... RSLNEIWQN. .MTWMEWERE IDNYTGLIYS LIEESQTQQE
D_CD_NDK_M  NSSWSN.... RSLDEIWQN. .MTWMEWERE IDNYTGLIYS LIEESQIQQE
D_UG_94UG1  NSSWSN.... RSVDEIWNN. .MTWMEWERE IDNYTELVYS LLEVSQIQQE
F1_BE_VI85  NSSWSN.... KSQEEIWNN. .MTWMEWEKE ISNYSNIIYK LIEESQNQQE
F1_BR_93BR  NSSWSN.... KSLEEIWGN. .MTWMEWEKE VSNYSKEIYR LIEDSQNQQE
F1_FI_FIN9  NSSWSN.... KSQDEIWNN. .MTWMQWEKE ISNYSKTIYM LIEKSQSQQE
F1_FR_MP41  NTSWSN.... KSHDEIWNN. .MTWMQWEKE INNYSNTIYR LIEESQNQQE
F2_CM_MP25  NLSWSN.... KSQDEIWGN. .MTWMEWEKE IGNYTDTIYR LIESAQNQQE
F2KU_BE_VI  NSSWSN.... RSQDEIWNN. .MTWMEWENE INNYTGIIYQ LIEQSQNQQE
G_BE_DRCBL  NTSWSN.... KSYNEIWEN. .MTWIEWERE IDNYTYHIYS LIEQSQIQQE
G_NG_92NG0  NTSWSN.... KSYNEIWDN. .MTWLEWERE IHNYTQHIYS LIEESQNQQE
G_SE_SE616  NVSWSN.... KSYNEIWDN. .MTWIEWERE INNYTYQIYS LLEESQNQQE
H_BE_VI991  NSSWSN.... KSLDEIWDN. .MTWMEWDKQ INNYTDEIYR LLEVSQNQQE
H_BE_VI997  NSTWSN.... KSLAEIWDN. .MTWMEWDRQ IDNYTEVIYR LLELSQTQQE
H_CF_90CF0  NSSWSN.... KSQSEIWDN. .MTWMEWDKQ ISNYTEEIYR LLEVSQTQQE
J_SE_SE702  NASWSN.... KSYEDIWEN. .MTWIQWERE INNYTGIIYS LIEEAQNQQE
J_SE_SE788  NASWSN.... KSYEDIWEN. .MTWIQWERE INNYTGIIYS LIEEAQNQQE
K_CD_EQTB1  NSSWSN.... KSQSEIWEN. .MTWMQWEKE ISNHTSTIYR LIEESQIQQE
K_CM_MP535  NSSWSN.... KSWEEIWNN. .MTWMEWEKE IGNYSDTIYK LIEESQTQQE
N_CM_YBF30  NETWSNN... TSYDTIWNN. .LTWQQWDEK VRNYSGVIFG LIEQAQEQQN
O_CM_ANT70  NRTWIG.... NES..IWDT. .LTWQEWDRQ ISNISSTIYE EIQKAQVQQE
O_CM_MVP51  NTSWSGRYN. DDS..IWDN. .LTWQQWDQH INNVSSIIYD EIQAAQDQQE
O_SN_99SE_  NTTWTNCTNT NKLDDIWDK. .LTWQQWDQQ ISNVSSIIYE EIRNAQVQQE
O_SN_99SE_  NRTWTN..NN TDLDTIWGN. .LTWQEWDQQ ISNISATIYD EIQKAQVQQE
U_CD___83C  NSSWSN.... KSLDNIWDN. .LTWMEWDRE ISNYTQVIYG LLEDSQKQQE
```

```
                751                                                   800
00BW0762_1  QNEKDLLALD SWKNLWSWFD ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW0768_2  RNERDLLALD SWKNLWSWFD ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW0874_2  KNEKDLLALD SWKNLWSWFD ISN.LWYIRI FIMIVGGLIG LRIVFAVLSL
00BW1471_2  KNEQELLALD SWENLWNWFS ISRWLWYIKI FIMIVGGLIG LRIIFAVLSV
00BW1616_2  KNEKDLLALD SWNSLWNWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW1686_8  QNEKDLLALD SWKNLWNWFN ISNWLWYIKI FIIIVGGLIG LRIIFAVLSI
00BW1759_3  INEKDLLALD SWKNLWNWFD ITKWLWYIKI FIMIVGGIIG LRIIFAVLSI
00BW1773_2  KNEKDLLALD SWKNLWNWFG ITKWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW1783_5  KNEKDLLALD SWNNLWNWFT ITNWLSYIKI FIMIVGGLIG LRIIFAVLSL
00BW1795_6  KNEKDLLALD SWKNLWNWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW1811_3  INEKDLLALD SWKNLWSWFD ISNWLWYIRI FIMIVGGLIG LRIIFAVLSI
00BW1859_5  KNEKDLLALD SWKNLWSWFD ITNWLWYIRI FIIIVGGLIG LRIIFAVLSI
00BW1880_2  QNEKNLLALD SWKNLN.WFS IT.HLWYIKI FIMIVGGLIG LRIVLVVLSV
00BW1921_1  KNEKDLLALD SWNNLWNWFS ITKWLWYIKI FIIIIGGLIG LRIIFAVLSI
00BW2036_1  QNEKDLLALD SWKNLWTWFD ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW2063_6  KNERDLLALD SWKNLWNWFN ISNWLWYIKI FIMIVGGLIG LRIIFAVLSM
00BW2087_2  NNEKDLLALD SWNNLWNWFD ITKWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW2127_2  QNEKDLLALD RWDSLWNWFG ISKWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW2128_3  KNEEDLLALD SWDSLWNWFS ITKWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW2276_7  QNEKDLLALD SWKNLWSWFD ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW3819_3  QNEKDLLALD SWKNLWTWFD ISNWLWYIKI FIMIVGGLIG LRIIFTVFSI
00BW3842_8  INEKDLLALD SWNSLWNWFD ITKWLWYIKI FIMVVGGLIG LRIIFAVLSI
00BW3871_3  KNEKDLLALD SWKNLWNWFD ITNWLWYIKI FIIIIGGLIG LRIIFAVLSI
00BW3876_9  QNEKDLLALD SWNSLWSWFD ITRWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW3886_8  KNEQELLALD SWKSLWNWFD ITNWLWYIKI FIMVVGGLIG LRIIFAVLSL
00BW3891_6  RNEKDLLAMD SWKNLWNWFD ITNWLWYIKI FIMIVGGLIG LRIVFAVLSI
00BW3970_2  QNEQDLLALN KWQHLWNWFD ITKWLWYIKI FIMIVGGLIG LRIIFAVLSI
00BW5031_1  QNEKDLLALD SWKNLWSWFD ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
 96BW01B21  KNEKDLLALD SWQNLWNWFS ITNWLWYIKI FIMIVGGLIG LRIVFAVLSI
  96BW0407  KNEKDLLALD SWNNLWNWFS ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
  96BW0502  KNEKDLLALD SWQNLWNWFS ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
 96BW06_J4  KNEKDLLALD SWKNLWNWFG ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
 96BW11_06  KDEKDLLALD SWQNLWNWFD IPKWLWYIKI FIMIVGGLIG LRIIFAVISM
  96BW1210  QNEKDLLALD SWANLWNWFN ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
 96BW15B03  QNEKDLLALD SWNNLWSWFN ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
 96BW16_26  KNEKDLLALD SWNSLWNWFS IVNWLRYIKI FIMIVGGLIG LRIIFAVLSI
 96BW17A09  KNEQELLALD SWANLWNWFA ISNWLWYIKI FIMIVGGLIG LRIVFAVLSV
 96BWMO1_5  RNEKDLLALD SWKTLWSWFD ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
 96BWMO3_2  KNEQDLLALD SWANLWNWFN ISNWLWYIKI FIMIVGGLIG LRIVFAVLSI
98BWMC12_2  ENEKDLLALD SWKNLWNWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
98BWMC13_4  KNEKDLLALD SWKNLWNWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSM
98BWMC14_a  QNEKDLLALD KWKDLRNWFN ISNWLWYIRI FIMIVGGLIG LRIIFAVLSI
98BWMO14_1  KNEKDLLALD SWNNLWTWFG ISSWLWYIKI FIMIVGGLIG LRIIFAVLSI
98BWMO18_d  RNEKDLLALD SWNNLWNWFN ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
98BWMO36_a  QNEKDLLALD SWQSLWNWFS ITKRLWYIKI FIMIVGGLIG LRIIFAVLSI
98BWMO37_d  QNEKDLLALD SWQNLWSWFS ITNWLWYIKI FIMIVGGLIG LRIIFAVLSL
99BW3932_1  KNEKDLLALD SWQNLWNWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
99BW4642_4  KNEKDLLALD SWKNLWTWFD ISNWLWYIRI FIMIVGGLIG LRIIFAVLSI
99BW4745_8  QNEKELLALD RWDTLWSWFN ITNWLWYIRL FIMIVGGLIG LRIIFAVLSI
99BW4754_7  KSEKDLLALD SWKNLWNWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSV
99BWMC16_8  RNEKDLLALD SWKNLWNWFN ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
A2_CD_97CD  KNEQDLLALD KWANLWNWFD ITRWLWYIKI FIMIVGGLIG LRIVIAIISV
A2_CY_94CY  KNEQDLLALD KWADLWSWFN ISHWLWYIRI FIMIVGGLIG LRIVFAIITV
A2D___97KR  KNEKDLLALD KWASLWNWFD ISNWLWYIKI FIMIVGGLIG LRIVMAIISV
A2G_CD_97C  KNKQDLLALD KWASLWNWFD ITKWLWYIKI FIMIVGGLIG LRIVFAVLSI
A_BY_97BL0  KNXQDLLALD KAG.LXSXXD ISNWLXYIXI FIIIVGGLIX LRIIFAVLSI
A_KE_Q23_A  KNEKELLELD KWANLWSWFD ISNWLWYIKI FIIIVGGLIG LRIVFAVLSV
A_SE_SE659  MNEQDLLALD KWANLWNWFD ITNWLWYIRI FIMIVGGLIG LRIVFAVLSI
A_SE_SE725  KNEQDLLALD KWANLWNWFE ITKWLWYIKI FIMIVGGLIG LRIVFAVFSV
A_SE_SE753  KNEQDLMALD KWTNLWTWFS ISNWLWYIRI FIMIVGGLIG LRIVFAVLAI
```

```
A_SE_SE853   KNEQDLLALD KWASLWNWFD ISRWLWYIRI FIMIVGGLIG LRIVFAVLSV
A_SE_SE889   KNEQDLLALD KWANLWNWFD ISKWLWYIKI FIMIVGGLIG LRIVIAVISI
A_SE_UGSE8   KNEKELLELD KWANLWNWFD ISNWLWYIKI FIMIVGGLIG LRIVIAVISI
A_UG_92UG0   RNEKDLLELD KWASLWNWFD ISKWLWYIKI FIMIVGGLIG LRIVFAVLTV
A_UG_U455_   KNELDLLALD KWANLN.WFN ISNWLWYIRL FVIIVGGLIG LRIVFAVLSV
AC_IN_2130   KNEQDLLALD KWADLWNWFD ISNWLWYIKI FIMIVGGLIG LRIVFTVLSI
AC_RW_92RW   KNEQDLLALD KWANLWNWFD ISNWLWYIKI FIMIVGGLIG LRIVFAVLSV
AC_SE_SE94   KNEQDLLALD KWASLWNWFD ISKWLWYIRI FIMIVGGLIG LRIIFAVLSI
ACD_SE_SE8   KNEQDLLALD KWGSLWNWFE ISKWLWYIRI FIMIVGGLIG LRIVFAVLSI
ACG_BE_VI1   QNEKDLLALD KWQNLWSWFN ISNWLWYIKI FIMIVGGLIG LRIVFAVLSV
AD_SE_SE69   KNEQDLLALD KWANLWSWFD ISNWLWYIKI FIMIVGGLIG LRIIFAILSI
AD_SE_SE71   MNEQDLLQLD KWASLWNWFD ITNWLWYIRI FIIIVGGLIG LRIVFAVLSL
ADHK_NO_97   XNEQDLLALD KWASLWNWFD ISKWLWYIKI FIMIVGGLIG LRIVFAVLSL
ADK_CD_MAL   KNEKELLELD KWASLWNWFS ISKWLWYIRI FIIVVGGLIG LRIIFAVLSL
AG_BE_VI11   KNEQDLLSLD KWASLWTWFD IANWLWYIRI FIMIVGGLIG LRVVFAVLNV
AG_NG_92NG   KNEQDLLALD KWASLWNWFD ITKWLWYIKI FIMIVGGLIG LRIVFAVLSI
AGHU_GA_VI   KNEQELLALD KWASLWSWFD ISKWLWYIRI FIMIVGGLIG LRIFFAVLSM
AGU_CD_Z32   INERDLLALD KWANLWNWFD ISNWLWYIKI FIMIVGGLIG LRIVFAVLSI
AJ_BW_BW21   KNEQDLLSLD KWASLWNWFS ISNWLWYIRI FIMIVGGLIG LRIIFAVLAI
B_AU_VH_AF   KNEQELLALD KWASLWNWFS ITKWLWYIKI FIMIVGGLVG LRIVFAVLSL
B_CN_RL42_   KNELELLELD KWASLWNWFD ITKWLWYIKI FIMIVGGLVG LRIVFAVLSI
B_DE_D31_U   KNEQELLELN KWENLWSWFD ISNWLWYIKI FIMIVGGLVG LRIVFAVLSI
B_DE_HAN_U   KNEQELLELD KWASLWSWYD ISNWLWYIKI FIMIVGGLIG LRIVFAVLSI
B_FR_HXB2_   KNEQELLELD KWASLWNWFN ITNWLWYIKL FIMIVGGLVG LRIVFAVLSI
B_GA_OYI__   KNEQELLELD KWAGLWSWFS ITNWLWYIRI FIIIVGGLVG LRIVFAVLSI
B_GB_CAM1_   KNEKDLLELD TWASLWNWFD ITNWLWYIKI FIMIIGGLIG LRIVFTILSL
B_GB_GB8_C   KNEQELLELD KWANLWNWFD ITNWLWYIKI FIMIIGGLIG LRIIFAVIST
B_GB_MANC_   KNEQELLELD KWGSLWSWFS ITNWLWYIKI FIMIVGGLVG LRIVFAVLSL
B_KR_WK_AF   KNEQELLELD KWASLN.WFN ITKWLWYIKI FIMIVGGLVG LRIIFFVLSI
B_NL_3202A   KNEQELLELD KWASLWNWFD ITKWLWYIKI FIMIVGGLVG LRIIFAVLSI
B_TW_TWCYS   KNEQDLLELD KWASLWNWFD ITNWLWYIRI FIMIVGGLIG LRIVFAVLSI
B_US_BC_L0   KNEQELLELD KWASLWNWFT ITNWLWYIKI FIMIVGGLIG LRIVFTVLSI
B_US_DH123   KNEQELLALD KWASLWNWFN ISNWLWYIKI FIMIVGGLIG LRIVFSVLSI
B_US_JRCSF   KNEQELLELD KWASLWNWFG ITKWLWYIKI FIMIVGGLIG LRIVFSVLSI
B_US_MNCG_   KNEQELLELD KWASLWNWFD ITNWLWYIKI FIMIVGGLVG LRIVFAVLSI
B_US_P896_   KNEKELLELD KWASLWNWFD ITNWLWYIRL FIMIVGGLIG LRIVFAVLSI
B_US_RF_M1   KNEQELLELD KWANLWNWFD ITQWLWYIRI FIMIVGGLVG LKIVFAVLSI
B_US_SF2_K   KNEQELLELD KWASLWNWFS ITNWLWYIKI FIMIVGGLVG LRIVFAVLSI
B_US_WEAU1   KNEQELLELD KWASLWTWFD ISNWLWYIKI FIMIVGGLIG LRIVFTVLSI
B_US_WR27_   KNEQELLELD KWASLWNWFN ITQWLWYIKI FXMIVGGLIG LRIVFAVLSI
B_US_YU2_M   KNEQELLALD KWASLWNWFD ITKWLWYIKI FIMIVGGLIG LRIVFVVLSI
BF1_BR_93B   KNEQELLALD KWASLWNWFD ISKWLWYIKI FIMIVGGLIG LRIVFAVLSI
C_BR_92BR0   KNEQDLLALD KWQNLWTWFG ITNWLWYIKI FIKIVGGLIG LRIIFAVLSI
C_BW_96BW0   KNEKDLLALD SWNNLWNWFG ITKWLWYIKI FIMIVGGLIG LRIIFAALSI
C_BW_96BW1   KNEKDLLALD SWKNLWNWFD ISKWLWYIKI FIMIVGGLIG LRIIFAVISM
C_BW_96BW1   QNEKDLLALD SWANLWNWFN ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
C_BW_96BW1   QNEKDLLALD SWNNLWSWFN ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
C_ET_ETH22   KNEKDLLALD KWENLWNWFN ITNWLWYIKI FIMIVGGVIG LRIIFAVLSI
C_IN_93IN1   KNEKDLLALD SWKNLWSWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
C_IN_93IN9   KNEKDLLALD SWKNLWSWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
C_IN_93IN9   RNEKDLLALD SWESLWNWFS ISKWLWYIKI FIMIVGGLIG LRIIFAVLSI
C_IN_94IN1   KNEKDLLALD SWKNLWSWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
C_IN_95IN2   ENEKDLLALD SWKNLWNWFD ITKWLWYIKI FIIIVGGLIG LRIIFAVISI
CRF01_AE_C   KNEQDLLALD KWANLWNWFS ITNWLWYIKI FIMVVGGLIG LRIVFAVLSI
CRF01_AE_C   RNEKDLLALD KWASLWNWFD ITRWLWYIKI FIIIVGGLIG LRIVFAVLSI
CRF01_AE_C   RNEKDLLELD KWTSLWNWFD ITRWLWYIRI FIMIVGGLIG LRIVFAVLSI
CRF01_AE_T   RNEKDLLELD KWASLWNWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
CRF01_AE_T   RNEKDLLELD KWANLWNWFS ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
CRF01_AE_T   RNEKDLLELD KWASLWNWFG ITKWLWYIKI FIMIVGGLIG LRIIFAVLSI
CRF01_AE_T   KNEKDLLELD KWASLWNWFD ITSWLWYIKI FIMIVGGLIG LRIIFAVLSI
CRF01_AE_T   RNEKDLLELD KWASLWNWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSI
```

```
CRF01_AE_T  RNEKDLLKLD KWASLWNWFD ISRWLWYIRI FIMIVGGLIG LRIVFAVLSI
CRF02_AG_F  KNEQDLLALD QWANLWNWFG ITKWLWYIKI FIMIVGGLIG LRIVFAVLTI
CRF02_AG_F  KNEKDLLALD KWESLWNWFN ITKWLWYIKI FIMIVGGLIG LRIVLAVLTV
CRF02_AG_G  KNEQDLLALD KWASLWNWFD ITKWLWYIKI FIMIVGGLIG LRIVFVVLAI
CRF02_AG_N  KNEQDLLALD KWASLWNWFD ITNWLWYIRI FIMIVGGLIG LRIVFAVLTI
CRF02_AG_S  KNEQDLLALD KWASLWDWFS ISSWLWYIRI FIIIVGGLIG LRIVFAVLAI
CRF02_AG_S  KNEQDLLALD KWASLWNWFD ITNWLWYILI FLMVVGGLIG LRIVFAVLAI
CRF03_AB_R  KNEQEILALD KWASLWNWFD ISKWLWYIKI FIMIVGGLVG LRIIFAVLSI
CRF03_AB_R  KNEQELLALD KWASLWNWFD ISKWLWYIKI FIMIVGGLVG LRIIFAVLSI
CRF04_cpx_  KNEQDLLALD KWASLWNWFS ITKWLWYIKI FIMIVGGLIG LRIIFAVLSI
CRF04_cpx_  KNEQDLLAFD KWANLWNWFN ISNWLWYIKI FIMIVGGLIG LRIIFAVLSI
CRF04_cpx_  KNEQDLLALD KWANLWSWFD ISHWLWYIKI FIMIVGGLIG LRIIFAVLSI
CRF05_DF_B  KNEQELLSLD QWASLWNWFD ITKWLWYIKI FIMIVGGLIG LRIVFAVLSI
CRF05_DF_B  KNEKDLLALD KWASLWNWFS ISNWLWYIRI FIMIVGGLIG LRIVFTVLSI
CRF06_cpx_  KNEQDLLALD KWANLWSWFD ITNWLWYIKI FIMIVGGLIG LRIVFTVLSV
CRF06_cpx_  KNEQDLLALD KWASLWSWFD ISNWLWYIRI FVIIVGGLLG LRIVFAVFSI
CRF06_cpx_  KNEQDLLALD KWASLWNWFD ISNWLWYIKI FIMIVGGLIG LRIVFAVLSI
CRF06_cpx_  KNEQELLALD KWASLWSWFD ISNWLWYIKI FIMIVGGLIG LKIVFAVLSL
CRF11_cpx_  KNEQDLLSLD KWASLWSWFD ITNWLWYIKI FIMIVGGLIG LRIIFAVLSV
CRF11_cpx_  KNEQDLLALD KWASLWNWFD ISNWLWYIRI FIMIVGGLIG LRIIFAVLSI
D_CD_84ZR0  KNEKELLELD KWASLWNWFS ITQWLWYIKI FIMIVGGLIG LRIVFAVLSI
D_CD_ELI_K  KNEKELLELD KWASLWNWFS ITQWLWYIKI FIMIIGGLIG LRIVFAVLSL
D_CD_NDK_M  KNEKELLELD KWASLWNWFS ITKWLWYIKL FIMIVGGLIG LRIVFAVLSV
D_UG_94UG1  KNEQELLKLD TWASLWNWFS ITQWLWYIKI FIMIVGGLIG LRIVFAVLSV
F1_BE_VI85  KNEQELLALD KWASLWNWFD ISNWLWYIKI FIMIVGGLIG LRIVFAVLSI
F1_BR_93BR  KNEQELLALD KWASLWNWFD ITQWLWYIKI FIMIVGGLIG LRIVFTVLSI
F1_FI_FIN9  RNEQELLELD KWDSLWSWFD ITNWLWYIKI FIMIVGGLIG LRIVFAVLSI
F1_FR_MP41  KNEQELLALD KWASLWSWFD ISNWLWYIKI FIMIVGGMIG LRIVFAVLSI
F2_CM_MP25  KNEQDLLALD KWDNLWNWFS ITRWLWYIEI FIMIIGSLIG LRIVFTVLSI
F2KU_BE_VI  KNEQDLLALD QWASLWSWFN ITQWLWYIKI FIMIVGGLIG LRIIFAVLSI
G_BE_DRCBL  KNEQDLLALD QWASLWSWFS ISNWLWYIRI FVMIVGGLIG LRIVFAVLSI
G_NG_92NG0  KNEQDLLALD KWASLWNWFD ISNWLWYIRI FIMIVGGLIG LRIVFAVLSI
G_SE_SE616  KNEQDLLALD QWASLWNWFG ITRWLWYIKI FIMIVGGLIG LRIVFAVLSI
H_BE_VI991  KNEQDLLALD KWANLWNWFS ITNWLWYIRI FIMIVGGIIG LRIVFAVLSI
H_BE_VI997  QNEQDLLALD KWDSLWNWFS ITNWLWYIKI FIIIVGALIG LRIIFAVLSI
H_CF_90CF0  KNEQDLLALD KWASLWTWFD ISHWLWYIKI FIMIVGGLIG LRIIFAVLSI
J_SE_SE702  TNEKDLLALD KWTNLWNWFN ISNWLWYIKI FIMIIGGLIG LRIIFAVLAI
J_SE_SE788  NNEKDLLALD KWTNLWNWFN ISNWLWYIKI FIMIIGGLIG LRIIFAVLAI
K_CD_EQTB1  KNEQDLLALD KWASLWNWFD ISNWLWYIKI FIMIVGGLIG LRIVFTVLSV
K_CM_MP535  KNEQDLLALD KWASLWNWFD ITKWLWYIKI FIMIIGGLIG LRIAFAVLSV
N_CM_YBF30  TNEKSLLELD QWDSLWSWFG ITKWLWYIKI AIMIVAGIVG IRIISIVITI
O_CM_ANT70  QNEKKLLELD EWASIWNWLD ITKWLWYIKI AIIIVGALVG VRVIMIVLNI
O_CM_MVP51  KNVKALLELD EWASLWNWFD ITKWLWYIKI AIIIVGALIG IRVIMIILNL
O_SN_99SE_  QNEKKLLELD EWASIWNWLD ITKWLWYIKI AIIIVGALIG VRIVMIVLNL
O_SN_99SE_  HNEKKLLELD EWASIWNWLD ITKWLWYIKI AIIIVGALIG VRIVMIVLNL
U_CD___83C  KSEKDLLELD KWASLWNWFD ITNWLWYIKI FIMIVGGLIG LRIVFTVFSI
```

```
            801                                                       850
00BW0762_1  VNRVRQGYSP LSFQTLTP.. NQR.GPDRLG GIEEEGGEQD RDRSIRLVSG
00BW0768_2  VNRVRQGYSP LSFQTLIP.. NPR.GPDRLG RIEEEGGEQD KDRSIRLVSG
00BW0874_2  VNRVRQGYSP LSFQTLTP.. SPR.EPDRLG RIEEEGGEQD KDRSIRLVGG
00BW1471_2  VNRVRQGYSP LSFQTLIP.. NPR.GPDRLE RIEEEGGEQD RGRSVRLVSG
00BW1616_2  VNRVRQGYSP LSFQTLTP.. NPR.ELDRLG RIEEEGGEQD RDRSIRLVSG
00BW1686_8  VNRVRQGYSP LSLQTLTP.. NPR.GPDRPR GIEEEGGEQD KDRSIRLVNG
00BW1759_3  VNRVRQGYSP SSFQTLIP.. NPE.GPDRLR RIEEEGGEQD RDRSIRLVNG
00BW1773_2  VNRVRQGYSP LSFQTLTP.. NPR.GPDRLG RIEEEGGEQD RDRSVRLVSG
00BW1783_5  VNRVRQGYSP LSFQTLIP.. NPR.GPDRLE RIEEEGGEQD RDRSIRLVSG
00BW1795_6  VNRVRQGYSP LSFQTLIP.. NPR.GPDRLG RIEEEGGEQD RDRSIRLVSG
00BW1811_3  VNRVRQGYSP LSFQTLIP.. NPG.GPDRLG RIEEEGGEQD RDRSVRLVNG
00BW1859_5  VNRVRQGYSP LSFQTLTP.. NPR.GPDRLG GIEEEGGEQD RDRSIRLVNG
00BW1880_2  VNRVRQGYSP LSLQTLSP.. NQR.GLDRLG RIEEEGGEQD RDKSIRLVSG
00BW1921_1  VNRVRQGYSP LSLQTLTP.. NPR.ELDRLG RIEEEGGEQD RGRSIRLVNG
00BW2036_1  VNRVRQGYSP LSFQTLTP.. NPR.GLDRLG RIEEEGGEQD RGRSIRLVQG
00BW2063_6  VNRVRQGYSP LSFQTLTP.. NPR.GPDKLE RIEEEGGEQD RNRSIRLVSG
00BW2087_2  VNRVRQGYSP LSFQTLTP.. NPR.EPDRLG RIEEEGGEQD RERSIRLVSG
00BW2127_2  VNRVRQGYSP LSFQTLTP.. NPR.GPDRLG RTEEEGGEQD NDRSIRLASG
00BW2128_3  VKRVRQGYSP LSFQTLTP.. NPG.GPDRLG RIEEEGGEQD REKSVRLVNG
00BW2276_7  VNRVRQGYSP LSFQTLTP.. NPR.GPDRLG GIEEEGGKQD RDRSIRLVSG
00BW3819_3  VNRVRQGYSP LSLQTLTP.. SPR.GPDGLR GIEEEGGEQD KDRSIRLVNG
00BW3842_8  VNRVRQGYSP LSLQTLTP.. NPR.ELDRLG RIEEEGGEQD RDRSIRLVDG
00BW3871_3  VNRVRQGYSP LSLQTLTP.. NPR.ELDRLG RIEEEGGEQD KDRSIRLVNG
00BW3876_9  VNRVRQGYSP LSFQTLTP.. NPR.EPDRPG RIEKEGGEQD KSRSIRLVSG
00BW3886_8  VNRVRQGYSP LSFQTLTP.. NTR.ELDRLG RIEEEGGEQG RDRSIRLLNG
00BW3891_6  VNRVRQGYSP LSFQTLIP.. NPR.GLDRLG RIEEEGGEQD RDRSIRLVNG
00BW3970_2  VNRVRQGYSP LSLQTLTP.. NQR.EPDRLG RIEEEGGEQD RKRSIRLVSG
00BW5031_1  VNRVRQGYSP LSFQTLTP.. SQR.ELDRLG RIEEEGGEQD RDRSIRLANG
 96BW01B21  VNRVRQGYSP LSFQTLTQ.. NPG.GPDGLG RIEEEGGEQD KDRSIRLVNG
  96BW0407  VNRVRQGYSP LSFQTLIP.. NPR.RSDRLG RIEEEGGEQD RGRSIRLVSG
  96BW0502  VNRVRQGYSP LPFQTLTP.. NPR.EPDRLG RIEEEGGEQD RGRSIRLVSG
 96BW06_J4  VNRVRQGYSP LSFQTLPP.. NPR.GIDRLG RIEEGGGEQD RDRSIRLVNG
 96BW11_06  VKRVRQGYSP LSFQTLTP.. NPR.GPDRLE RIEEEGGEQD RNRSIRLVSG
  96BW1210  VNRVRQGYSP LSFQTLIP.. SPR.EPDKLE RIEEEGGEQD RDRSIRLVSG
 96BW15B03  VNRVRQGYSP LSFQTLTP.. NPR.GLDRLG RIEEEGGEQD RDRSIRLVQG
 96BW16_26  VNRVRQGYSP LSFQTLTP.. NPR.GPDRLG RIEEEGGEQD RDRSVRLVHG
 96BW17A09  VNRVRQGYSP LSLQTLTP.. NPR.GPDGLE RIEEGGGEQD RGRSIRLVSG
 96BWMO1_5  VNRVRQGYSP LSFQTLTP.. NPR.GPDRLE GIEEEGGEQD KDRSIRLVSG
 96BWMO3_2  VNRVRQGYSP LPFQTLTP.. NPR.ELDRLG RIEEEGGEQD RDRSIRLVSG
98BWMC12_2  VNRVRQGYSP LSFQTLNP.. NPR.GLDRLG RIEEEGGEQD KSTSIRLVNG
98BWMC13_4  VNRVRQGYSP LSFQTLTP.. NPR.GPDRLE RIEEEGGEQD RDRSIRLVNG
98BWMC14_a  VNRVRQGYSL TSLQTHLP.. NAG.GLDRLD RIGEEGGEQD RHRSIRLVSG
98BWMO14_1  VNRVRQGYSP LSFQTLTP.. VPR.EPDRLG GIEEEGGEQD RDRSVRLVNG
98BWMO18_d  VNRVRQGYSP LSFQTLTP.. NPR.GPDRLG RIEEEGGEQD RDRSIRLVSG
98BWMO36_a  INRVRQGYSP LSFQTLTP.. TPR.DPDRLR GIEEEGGEQD RDRSIRLVSG
98BWMO37_d  VNRVRQGYSL TSFQTVIP.. NPR.GPDRPR GIEEEGGEQD RDRSIRLVSG
99BW3932_1  VNRVRQGYSL LSFQTLTP.. NPR.GPDRLG GIEEEGGEQD RDRSIRLVNG
99BW4642_4  VNRVRQGYSP LSFQTLTP.. NPR.ELDRLG RIEEEGGEQD RDRSVRLVNG
99BW4745_8  VNRVRQGYSP LSLQTLTP.. SPR.RPDRLG GIEEEGGEQD RTRSVRLVNG
99BW4754_7  VNRVRQGYSP LSFQTLTP.. NQR.GPDRLG EIEEEGGEQD RDRSIRLVNG
99BWMC16_8  VNRVRQGYSP LSFQTLAP.. NPG.GLDRLG RIEEEGGEQD RGRSIRLVNG
A2_CD_97CD  VKRVRQGYSP LSFQIPTP.. NPE.GLDRPG RIEEEGGEQG RDRSIRLVSG
A2_CY_94CY  VNRVRQGYSP VSFQIPTP.. SPE.GPDRPR GTEEGGGEQG RDRSIRLVNG
A2D___97KR  VNRVRQGYSP VSFQIPPP.. TPE.DPDRHG RIEDGGGEQG RDRSVRLVSG
A2G_CD_97C  VNRVRQGYSP LSFQTLTH.. HQR.EPDRPE RIEEGGGEQD RDRSVRLVSG
A_BY_97BL0  INRARXXYSP LSLQTLTP.. HPE.RPDRPX RIKEEGXEQG RDRSIRLVSG
A_KE_Q23_A  INRVRQGYSP LSFQTHTP.. NPR.GLDRPE RIEEEDGEQG RGRSIRLVSG
A_SE_SE659  INRVRQGYSP LSFQTHTP.. NPG.GLDRPG RIEEEGGEQG RDRSIRLVSG
A_SE_SE725  INRVRQGYSP LSFQTHTP.. DPR.GLDRPR RIEEEGGEQG RGRSIRLVSG
A_SE_SE753  INRVRQGYSP LSFQIHTP.. NPR.DPDRPG RIEEEGGEQG RDRSIRLASG
```

```
A_SE_SE853   INRVRQGYSP LSFQIHTP.. NPG.DLDRPG RIEEEGGEQD RGRSIRLVSG
A_SE_SE889   INRVRQGYSP LSFQIHTP.. SPG.GLDRPG RIEEEGGEQD RNRSIRLVNG
A_SE_UGSE8   IKRVRQGYSP LSFQIHTP.. SPR.DPDRPG RIEEEGGEQG RDRSIRLVSG
A_UG_92UG0   INRVRQGYSP LSFQTHTP.. NPR.GLDRPG RIEEEGGEQD RGRSIRLVSG
A_UG_U455_   INRVRQGYSP LSFQTLAP.. IPE.GLGRPG RIEEEGGEQG KDRSIRLVSG
AC_IN_2130   INRVRQGYSP LSFRTHTP.. NPG.GLDRPG RIEEEDGEQG KDRSIRLVSG
AC_RW_92RW   VNRVRQGYSP LSFQTLIP.. NPR.GPDRLG GIEEEGGEQD RGRSIRLVSG
AC_SE_SE94   INRVRQGYSP LSFQIHTP.. NPG.GPDRPG RTEEEGGEQD RDRSIRLVSG
ACD_SE_SE8   INRVRQGYSP LSFQTHTP.. NPE.GVDRPG RIEEEGGEQG RDRSIRLVSG
ACG_BE_VI1   VNRVRQGYSP LSFQTLIP.. NPR.GPDRPG EIEEGGGEQD RDTSTRLVSG
AD_SE_SE69   VNRVRQGYSP LSFQTLLP.. APRGP.DRPE GIEEEGGEQG RGRSIRLVNG
AD_SE_SE71   INRVRQGYSP LSFQTHTP.. NPR.DLDRPE RIEEGGGEQD RTRSIRLVSG
ADHK_NO_97   VNRVRQGYSP LSFQTLIP.. NPRGA.DRPE GIEEEGGEQD XNRSIRLVNG
ADK_CD_MAL   VNRVRQGYSP LSLQTLLP.. TPRGPPDRPE GIEEEGGEQG RGRSIRLVNG
AG_BE_VI11   INRVRQGYSP LSFQILTP.. HQRDP.DRPG RIEEGGGEQD RDTSRRLVGG
AG_NG_92NG   VNRVRQGYSP LSFQTLTH.. HQREP.DRPE RIEEGGGEQD RDRSVRLVSG
AGHU_GA_VI   VNRVRQGYSP LSFQTLFP.. NQREP.DRPE GIEEEGGEQG RSRSIRLVNG
AGU_CD_Z32   INRVRQGYSP LSFQTLTH.. HQR.EPDRPE GIEEGGGEQD RDRSIRLVSG
AJ_BW_BW21   VNRVRQGYSP LSLQTLIP.. NPTEV.DRPG GIEEGGGEQG KTRSIRLVNG
B_AU_VH_AF   VKKVREGYSP LSLQIRPP.. APRGP.DRPE GIEEEGGEQD RDKSVRLVDG
B_CN_RL42_   VNRVRQGYSP LSLQTRFP.. AQRGP.GRPE GIEEEGGERD RDRSERLVTG
B_DE_D31_U   VNSVRQGYSP LSFQTRLP.. APRGP.DRPE GIEEEGGDRD RDRSNRLVKG
B_DE_HAN_U   VNRVRQGYSP LSFQTLLP.. ATRGP.RQPE EIEEEGGERD RGRSVRLVSG
B_FR_HXB2_   VNRVRQGYSP LSFQTHLP.. TPRGP.DRPE GIEEEGGERD RDRSIRLVNG
B_GA_OYI__   VNRVRQGYSP LSFQTRLP.. TQRGP.DRPE GIEEEGGERD RDRSGRLVDG
B_GB_CAM1_   VNRVRQGYSP LSFQTRFP.. VPRGP.DRPE GIEEEGGGRD RDTSGRLVTG
B_GB_GB8_C   VNRVRQGYSP LSLQTHLP.. TPRGP.DRPE GIEEEGGEQD RDRSIRLVNG
B_GB_MANC_   VNRVRQGYPP LSFQTHLP.. APRGP.DRPG GIEEEGGEGD RDRSSRLVHG
B_KR_WK_AF   VNRVRQGYSP LSFQTHFP.. SPRGP.DRPG GIEGEGGEGD SGGSSHLVDG
B_NL_3202A   VNRVRQGYSP LSFQTRLP.. APRGP.DRPE GIEEEGGERD RDRSGRLVNG
B_TW_TWCYS   VNRVRQGYSP LSFQTHLP.. TPRGP.DRPE GIDEEGGERD RDRSERSVDG
B_US_BC_L0   VNRVRQGYSP LSFQTHLP.. TPRGP.DRPE GIEEEGGERD RDRSGRLVNG
B_US_DH123   VNRVRQGYSP LSFQTRFP.. ASRGP.DRPE GIEEEGGDRD RDRSSPLVDG
B_US_JRCSF   VNRVRQGYSP LSFQTLLP.. ATRGP.DRPE GIEEEGGERD RDRSGQLVNG
B_US_MNCG_   VNRVRQGYSP LSLQTRPP.. VPRGP.DRPE GIEEEGGERD RDTSGRLVHG
B_US_P896_   VNRVRQGYSP LSFQTLLP.. ASRGP.DRPE GTEEEGGERD RDRSGPLVNG
B_US_RF_M1   VNRVRQGYSP LSFQTHLP.. APRGP.DRPE GIEGEGGERD RDRSGGAVNG
B_US_SF2_K   VNRVRQGYSP LSFQTRLP.. VPRGP.DRPD GIEEEGGERD RDRSVRLVDG
B_US_WEAU1   VNRVRQGYSP LSFQTHLP.. APRGP.DRPE GIEEEGGERD RDRSGRLVDG
B_US_WR27_   XNRVRQGXSP LSFQTLLP.. VPRGP.DRPE GIEEEGGERD RDRSNRLVHG
B_US_YU2_M   VNRVRQGYSP LSFQTHLP.. AQRGP.DRPD GIEEEGGERD RDRSGPLVDG
BF1_BR_93B   VNRVRKGYSP LSLQTRFP.. SPREP.DRPE GIEEGGGEPG KDRSVRLVNG
C_BR_92BR0   VNRVRQGYSP LSFQTLTP.. NPR.GPDRLG GIEEEGGEQD RDRSIRLVSG
C_BW_96BW0   VNRVRQGYSP LSFQTLTP.. NPR.GPDRLG GIEEEGGEQD RDRSIRLVSG
C_BW_96BW1   VNRVRQGYSP LSFQTLTP.. NPR.GPDRLE RIEEEGGEQD RNRSIRLVSG
C_BW_96BW1   VNRVRQGYSP LSFQTLIP.. SPR.EPDKLE RIEEEGGEQD RDRSIRLVSG
C_BW_96BW1   VNRVRQGYSP LSFQTLTP.. NPR.GLDRLG RIEEEGGEQD RDRSIRLVQG
C_ET_ETH22   VNRVRQGYSP LSFQTLIP.. HPR.GPDRLG GIEEEGGEQG RDRSIRLVNG
C_IN_93IN1   VNRVRQGYSP LSFQTLTP.. NPR.GPDRLG RIEEEGGEQD KDRSIRLVNG
C_IN_93IN9   VNRVRQGYSP LSFQTPTP.. NPG.GPDRLG RIEEEGGEQG KDRSIRLVNG
C_IN_93IN9   VNRVRQGYSP LSFQTLTP.. NPG.GPDRLG RIEEEGGEQD KNRSIRLVNG
C_IN_94IN1   VNRVRQGYSP LSFQTPTP.. NPG.GPDRLG RIEEEGGGQD NVRSIRLVNG
C_IN_95IN2   VNRVRQGYSP LSFQTLTP.. NPG.GPDRLG RIEEEGGEQD KDRSIRLVSG
CRF01_AE_C   VNRVRQGYSP LSFQTLTH.. HQR.EPDRPE RIEEGGGEEG KDRSIRLVSG
CRF01_AE_C   VNRVRQGYSP LSFQTLTH.. QQR.EPDRPE RIEEESGEQG RDRSIRLVSG
CRF01_AE_C   VNRVRQGYSP LSFQTLTH.. HQR.EPDRPE RIEEGGGEQD KDRSVRLVSG
CRF01_AE_T   VNRVRQGYSP LSFQTPIH.. HQR.EPDRPE RIEEGGGEQG RDRSVRLVSG
CRF01_AE_T   VNRVRQGYSP LSLQTPTH.. HQR.EPDRPE RIEEGGGEQG RDRSVRLVSG
CRF01_AE_T   VNRVRQGYSP LSFQTPTH.. HQRGEPDRPE RIEEGGGEQG RDRSVRLVSG
CRF01_AE_T   VNRVRQGYSP LSFQTPLH.. HQR.EPDRPE RIEEEGGEQG RDRSVRLVSG
CRF01_AE_T   VNRVRQGYSP LSFQTPSH.. HQK.EPDRPE GIEEGGGEQG RDRSVRLVSG
```

```
CRF01_AE_T  VNRVRQGYSP LSFQTLSH.. HQR.DPDRPE RIEEGGGEQG RDRSVRLVSG
CRF02_AG_F  INRVRQGYSP LSFQTLTH.. HQR.GPDRPE RIEEGGGEQD RDRSGRLVSG
CRF02_AG_F  IKRVRQGYSP LSFQTLTH.. HQR.EPDRPE RIEEGGGEQD KDRSVRLVSG
CRF02_AG_G  INRVRQGYSP LSFQILTP.. NPR.GPDRPE GIEEGGGEQD RDRSIRLVSG
CRF02_AG_N  INRVRQGYSP LSFQTLTH.. HQR.EPDRPE RIEEGGGEQD KDRSVRLVSG
CRF02_AG_S  INRVRQGYSP LSFQTLTH.. HQR.EPGRPE RIEEGGGGQD RDRSVRLVSG
CRF02_AG_S  INRVRQGYSP LSFQTLTH.. HQR.GADRPE GIEEGGGEQD RDRSVRLVSG
CRF03_AB_R  VNRGRQGYSP LSFQTRLP.. AQRGP.DRPE GIEEEDGERD RDTSIRLVNR
CRF03_AB_R  VNRVRQGYSP LSFQTRLP.. TQRGP.DRPE GIEEEGGERD RDTSIRLVNG
CRF04_cpx_  VNRVRQGYSP LSLQTLIPT. TQRGL.DRPG GTEEEGGEQD RSRSIRLVNG
CRF04_cpx_  VNRVRQGYSP LSFQTLIP.. TQREP.GRPE GTEEEGGEQD RSRSIRLVNG
CRF04_cpx_  VKRVRQGYSP LSSQTLIPT. TQRGP.DRPE GTEGGGGEQD RIESIRLVNG
CRF05_DF_B  VNRVRKGYSP LSFQTPFP.. APRGP.DRPE ETEEGGGEQD RGRSIRLVNG
CRF05_DF_B  VNRVRKGYSP LSFQTLLP.. GPRGP.DRPE GTEEEGGEQG RDRSVRLLTG
CRF06_cpx_  VNRVRQGYSP LSLQTLIP.. NPTGA.DRPG EIEEGGGEQG RTRSIRLVNG
CRF06_cpx_  VNRVRQGYSP LSLQTLIP.. NPTGA.DRPG EIEEEGGEQG RTRSIRLVNG
CRF06_cpx_  VNRVRQGYSP LSLQTLIP.. NPAEV.DRPG GIEEGGGEQG RNRSIRLVNG
CRF06_cpx_  VKRVRQGYSP LSLQTLIP.. NSAGV.DRPG EIEGGGGEQD RTRSIRLVNG
CRF11_cpx_  VNRCRQGYSP LSFQTLNP.. TQQEA.DRPG GIEEGGGEQG RTRSIRLVSG
CRF11_cpx_  VNRCRQGYSP LSFQALTP.. SQQEP.DRPG GTKEGGGEQG RTSSIRLVSG
D_CD_84ZR0  VNRVRQGYSP LSFQTLLP.. APRGP.DRPE GIEEEGGEQD RGTSIRLVNG
D_CD_ELI_K  VNRVRQGYSP LSFQTLLP.. APRGP.DRPE GTEEEGGERG RDRSVRLLNG
D_CD_NDK_M  VNRVRQGYSP LSFQTLLP.. VPRGP.DRPE EIEEEGGERG RDRSIRLVNG
D_UG_94UG1  VNRVRQGYSP LSFQTLLP.. APREP.DRPE GIEEEGGERD RGRSIRLVNG
F1_BE_VI85  VNRVRKGYSP LSLQTLIP.. SPRGP.DRPE GIEEGGGEQG KDRSVRLVTG
F1_BR_93BR  VNRVRKGYSP LSFQTHIP.. SPREP.DRPE GIEEGGGEQG KDRSVRLVTG
F1_FI_FIN9  VNRVRKGYSP LSLQTLIP.. APTEP.DRPE GIEEGGGEQG KDRSVRLVNG
F1_FR_MP41  VNRVRKGYSP LSLQTLIP.. SPRGP.DRPE GIEEGGGEQD RNRSVRLVNG
F2_CM_MP25  INRVRQGYSP LSLQTLIP.. NSRGP.ERPG GIEEEGGEQD KDRSIRLVSG
F2KU_BE_VI  VNRVRQGYSP LSFQTLIP.. SPRGP.DRPE GTEERGGEQD RDRSTRLVSG
G_BE_DRCBL  VNRVRQGYSP LSFQTLLH.. HQREP.DRPA GIEEGGGEQD RDRSIRLVSG
G_NG_92NG0  VNRVRQGYSP LSFQTLTH.. HQREP.DRLG KTEEGGGEQD RDRSTRLVSG
G_SE_SE616  VNRVRQGYSP LSFQTLTH.. HQREP.DRPE GIEEGGGEQG RGRSVRLVSG
H_BE_VI991  VNRVRQGYSP LSLQTLIP.. NQRGP.DRPR EIEEEGGEQD RDRSIRLVNG
H_BE_VI997  VGRVRQGYSP LSFQTLIP.. NPRGP.DRPE GIEEEGGEQD RGRSVRLVNG
H_CF_90CF0  VNRVRQGYSP LSFQTLVP.. NPRGP.DRPE GTEEGGGEQD RDRSVRLVNG
J_SE_SE702  VNRVRQGYSP LSFQTLIP.. NPTEA.DRPG GIEEGGGEQG RTRSIRLVNG
J_SE_SE788  VNRVRQGYSP LSFQTLIP.. NPTEA.DRPG GIEEGGGEQG RTRSIRLVNG
K_CD_EQTB1  VNRVRQGYSP LSFQTLTP.. SPRGP.DRPE GIEEGGGEQD KDRSVRLVSG
K_CM_MP535  VNRVRQGYSP LSFQTLIP.. TSRGA.DRPE GIEEEGGEQD KNRSVRLVSG
N_CM_YBF30  IARVRQGYSP LSLQTLIPT. ARG..PDRPE ETEGGVGEQD RGRSVRLVSG
O_CM_ANT70  VKNIRQGYQP LSLQIPNHH. QEE..AGTPG RTGGGGGEEG RPRWIPSPQG
O_CM_MVP51  VKNIRQGYQP LSLQIPVPH. RQE..AETPG RTGEEGGEGD RPKWTALPPG
O_SN_99SE_  VKNIRQGYQP LSFKTHIHH. QPE..AEAPG GTGEGGGERG MPTLIPWPQG
O_SN_99SE_  VRNIRHGYQP LSFQTPTHHQ QPE..AQAPG GTGEGGGERD RLRSIPSPQG
U_CD___83C  INRVRQGYSP LSFQTLLP.. TPR.GPDRPG RTEEEGGEED NNRSVRLVNG
```

```
            851                                                   900
00BW0762_1  FLALFWDDLR SLCLFSYHRL RDFILVAARA VELLGRSSLK GLQRGWEILK
00BW0768_2  FLALVWDDLR RLCLFSYHRL RDFILVATRA VELLGHSSLR GLQRGWEALK
00BW0874_2  FLALAWDDLR SLCLFSYHRL RDFILIAART VELLGRSSLR GLQRGWEILK
00BW1471_2  FLTLAWDDLR SLCIFLYRLL SDFISIAART VNLLGQNSLR GLQRGWEALK
00BW1616_2  FLALAWDDLR SLCLFSYHRL RDFTLIAARA VETLGR.... ...RGWEALK
00BW1686_8  FLALAWDDLR SLCLFSYHQL RDFILIVARA VELLGRNSLR GLQRGWEILK
00BW1759_3  FLALFWDDLR SLCLFSYHRL RDLILVTARA VELLG..... ..QRGWEALK
00BW1773_2  FLALTWDDLR SLCLFCYHRL RDFILIAARV VELLGRSSLR GLQKGWEALK
00BW1783_5  FLALAWDDLR SLCLFSYHRL RDFILVTARA VELLGRSSLR GLQRGWEILK
00BW1795_6  FLALAWDDLR SLCLFCYRRL RDFILVTARA VELLGRSSLK GLQRGWEILK
00BW1811_3  FFALAWDDLR SLCLFCYHRL RDFILVTARA VELLGHSSLK GLQRGWEILK
00BW1859_5  FLALAWDDLR SLCLFSYHRL RDCILIAARA VELLGHSSLR GLQRGWEVLK
00BW1880_2  FLALAWDDLR SLCLFSYHRL RDFILVTARA VELLGRSGLK GLQRGWEILK
00BW1921_1  FLPLVWEDLR SLCLFSYHRL RDLILVVARA VELLGRSSLR GLQKGWETLK
00BW2036_1  FLALAWDDLR NLCLFSYHQL RDLILVTTRV VELLGRSSLR GLQRGWEALK
00BW2063_6  FLALAWDDLR SLCLFCYHRL KDFVLVTARV VELLGLSSLK GLQRGWEILK
00BW2087_2  FLPLFWDDLR SLCLFSYHRL RDLILIAARA VELLGRSSLR GLQRGWEILK
00BW2127_2  FLAPAWDDLR SLCLFSYHRL RDLILVTARV VELLGRS... ....GWEALK
00BW2128_3  FLALFWDDLR SLCLFSYHRL RDFILIAARV VELLGRSSLR GLQRGWETLK
00BW2276_7  FLALAWDDLR SLCLFSYHRL RDFILVTARA VELLGRSSLR GLQRGWEVLK
00BW3819_3  FLALAWDDLR SLCLFSYHRL RDLTLVTARG IELMGRSSLR GLQKGWEALK
00BW3842_8  FLALAWEDLR SLCLFSYHRL RDLILVTARA VELLGRSSLR GLQRGWEALK
00BW3871_3  FLALAWDDLR SLCLFSYHRL RDFILIVARV VELLGRSSLR GLQKGWETLK
00BW3876_9  FLALAWDDPR SLCLFSYHRL RDFILVVVRA VELLGRSSLK GLKRGWEALE
00BW3886_8  FLAIAWDDLR SLCLFSYHRL RDFILLIARA VELLGRSSLK GLQRGWEALK
00BW3891_6  FLALAWEDLR SLCLFSYHRL RDFILVTARA VELLGRSSLR GLQRGWEALK
00BW3970_2  FLALAWDDLR SLCLFSYHHL RDFILIAARV VELLGR.... ...RGWDILK
00BW5031_1  FLALAWEDLR SLCLFSYRHL RDFILIVVRA VELLGRSSLR GIQKGWDALK
 96BW01B21  FLPLVWDDLR NLCLFSYHRL RDFILVIARA VELLGRSSLR GLQRGWETLK
  96BW0407  FLALAWDDLR SLCLFSYHRL RDFILIAARA AELLGRSGLR GLQKGWETLK
  96BW0502  FLALAWDDLR SLCLFSYHRL RDFILIAARV LELLG..... ..QRGWEALK
 96BW06_J4  FLALAWDDLR SLCLFSYHQL RDFILVVARA VELLGRSSLR GLQRGWEALK
 96BW11_06  FLALAWDDLR SLCLFCYHRL RDFTLVTARA VELLGRSSLK GLQRGWEILK
  96BW1210  FLALAWDDLR SLCLFSYHRL RDSILVAART VELLGRSSLR GLQRGWEALK
 96BW15B03  FLALAWDDLR SLCLFSYHRL RDLILVTARV VELLGRSSLR GLQRGWEALK
 96BW16_26  FLALAWDDLR SLCLFSYHRL RDFILVAVRV VELLGR.... ...RGWEALK
 96BW17A09  FLALAWDDLR SLCIFLYHHL RDFILIAART VNLLGQSSLR GLQRGWEALK
 96BWMO1_5  FLALAWDDLR SLCLFCYHRL RDFILVTARA VELLGRSSLK GLQRGWEILK
 96BWMO3_2  FLALAWDDLR SLCLFSYHRL RDFLLVTVRA AELLGRSSLR GLQRGWEALK
98BWMC12_2  FLAIAWDDLR SLCLFSYHRL RDFILIAARA VELLGRSSLR GLQRGWEALK
98BWMC13_4  FLALAWDDLR NLCLFCYHRL RDFILVTARA VELLGRSSLT GLQRGWEILK
98BWMC14_a  LLALAWDDLR SVRLFSYHQL RNFILIVARA VELLGR.... ...RGWETLK
98BWMO14_1  FLALFWDDLR SLCLFSYHRL RDLILIAVRA VELLGRSSLW GLQKGWEALK
98BWMO18_d  FLALAWDDLR SLCLFSYHRL RDFILIAARA VELLGHSILR GLQRGWEILK
98BWMO36_a  FLALAWDDLR SLCLFSYHRL RDFILVTARV VELMGRSSLK GLQRVWEILK
98BWMO37_d  FLALAWDDLR SLCLFSYHQL RDFILLIARV VERLGYSSLR GLQRGWEALK
99BW3932_1  FFSLAWDDLR SLCLFSYHRL RDLILVTVRV VELMGRCSLR GLQRGWEALK
99BW4642_4  FLALAWDDLR SLCLFSCHRL RDCILIAVRA VELLGRSSLK GLQRGWEALK
99BW4745_8  FLALAWDDLR SLCLFSYHHL RDFILVTARA VELLGRSSLR GLQRVWEALK
99BW4754_7  FLAIVWDDLR SLCLFSYHRL RDFILIATRA VELLGRSSLK GLQRGWEILK
99BWMC16_8  FLAIAWEDLR SLCLFSYHRL RDLTLIVTRA VELLGRSSLK GLQRGWEALK
A2_CD_97CD  FLALAWDDLR SLCLFSYHRL RDCILIAARI VELVGHSSLK GLRLGWEGLK
A2_CY_94CY  FFALAWDDLR SLCLFSYHRL RDCILIAART VELLGHCSLK GLRLGWEGLK
A2D___97KR  FLALVWEDLR SLCLFSYHRL RDCISIAART VELLGHSSLK GLRLGWEGLK
A2G_CD_97C  FLALAWDDLR SLCLFCYHRL RDSILIAART VELLRHSSLK GLRLGWEGLK
A_BY_97BL0  FLALAXDDLX SLCLFSYHRL RDFISIAART XELLKRSSLX GLRLXXXGLK
A_KE_Q23_A  FLALAWDDLR SLCLFSYHRL RDFILIAART VELLGHSSLK GLRLGWEGIK
A_SE_SE659  FLALAWDDLR SLCLFSYHHL RDLILIAART VELLGHSSLK GLRLGWEGLK
A_SE_SE725  FLALAWDDLR SLCLFSYHRL RHFILIATTT VELLGHSSLK GLRLGWEGLK
A_SE_SE753  FLTLAWEDLR SLCLFSYHRL RDLILIAART VELLGHSSLK GLRLGWEGLK
```

```
A_SE_SE853   FLALAWDDLR SLCLFSYHRL RDFILIAART VELLGQR... ....GWEGLK
A_SE_SE889   FLALAWDDLR SLCLFCYRRL RDFILIVART VELLGHSSLR GLRLGWEGLK
A_SE_UGSE8   FLALAWDDLR SLCRFSYHRL RDFISIATRI VELLG..... ......QGLK
A_UG_92UG0   FLALAWDDLR NLCLFSYHRL RDFILIAART VELPGHSSLK GLRLGWEGLK
A_UG_U455_   FLAIAWDDLR NLCLFSYHRL RDFALIVARA VELLGRSSLK GLRLGWEGLK
AC_IN_2130   FLALFWDDLR SLCLFSYHRL RDFILIAKRT VELLGHSSLK GLRLGWEGLK
AC_RW_92RW   FLALAWDDLR SLCLFSYHRL RDLLLIAART VELLGRSSLR GLQRGWETLK
AC_SE_SE94   FLALAWDDLR SLCLFSYHRL RDFILI.... ..LG.HSSLK GLRLGWEGLK
ACD_SE_SE8   FLALAWDDLR SLCLFSYHRL RDLILIAARI VELLGRR... ....GWEAIK
ACG_BE_VI1   FFALAWDDLR SLCIFLYHHL RDLLLIATRA VELLG..... ..QRGWEALK
AD_SE_SE69   FSALIWDDLR NLCLFSYHRL RDLLLIALRI VELLGRR... ....GWEALK
AD_SE_SE71   FLAIAWDDLR SLCLFSYHRL RDLILIAARI VELLGRR... ....GWEALK
ADHK_NO_97   FLPLVWEDLR NLCLFSYRLL RDLLLILART VTLLGSR... ....GWETLK
ADK_CD_MAL   FSALIWDDLR NLCLFSYHRL RDLLLIATRI VELLGRR... ....GWEALK
AG_BE_VI11   FLTLVWDDLR SLCLFSYHRL RDLVLIAART LELLGRSGLR GLRLGWEGLK
AG_NG_92NG   FLALAWDDLR NLCLFSYHRL RDLVLIAART AELLRRSSLQ GLRLGWEGLK
AGHU_GA_VI   FLPLIWEDLR NLCLFSYRHL RDLLLIVART VELLGKR... ....GWGALK
AGU_CD_Z32   FLPLAWDDLR SLCLFCYHRL RDCALIAARI VETLIRR... ....GWETLK
AJ_BW_BW21   FLALAWEDLR NLCLFSCHRL RDFALIAART VDTLGRR... ....GWEILK
B_AU_VH_AF   FLALIWDDLR SLCLFSYHRL RDLLLIAARI VELLGRR... ....GWEALK
B_CN_RL42_   FLSLIWEDLR SLCLFSYHRL RDLLLIVARI VELLGRR... ....GWEVLR
B_DE_D31_U   FLALIWDDLS SLCLFLYHRL RDLLLIAARI VELLGRR... ....GWEVLK
B_DE_HAN_U   FLALFWDDLR SLCLFSYRRL RDLLLIVAKI VETLGRR... ....GWEVLK
B_FR_HXB2_   SLALIWDDLR SLCLFSYHRL RDLLLIVTRI VELLGRR... ....GWEALK
B_GA_OYI__   FLALIWDDLR SLCLFSYHRL RDLILIVARI VELLGRR... ....GWEVLK
B_GB_CAM1_   FLALIWDDLR SLCLFSYHRL RDLLLIVARI VELLGRR... ....GWEALK
B_GB_GB8_C   FLALFWDDLR SLCLFSYHRL RDLLLIVTRI VGLLGRR... ....GWEALK
B_GB_MANC_   FLALVWDDLR SLCLFSYHRL RDLLSIAARI VELLGRR... ....GWEILK
B_KR_WK_AF   FLTLIWVDLR SLCLFSYHLL RDLLLIVTRS VELLGLR... ....GWEILK
B_NL_3202A   FLALIWDDLR SLCLFSYHRL RDLLLIVTRI VELLGRR... ....GWEVLK
B_TW_TWCYS   FLAIIWVDLR SLCLFSYHRL RDLLLIVTRI VELLGRR... ....GWEALK
B_US_BC_L0   FLALFWDDLR SLCLFSYHRL RDLILVVTRI VELLGRR... ....GWEALK
B_US_DH123   FLAIIWVDLR TLFLFSYHRL RDLLLIVTRI VELLGRR... ....GWELLK
B_US_JRCSF   FLALIWVDLR SLFLFSYHRL RDLLLTVTRI VELLGRR... ....GWEILK
B_US_MNCG_   FLAIIWVDLR SLFLFSYH.H RDLLLIAARI VELLGRR... ....GWEVLK
B_US_P896_   FLALFWVDLR NLCLFLYHLL RNLLLIVTRI VELLGRR... ....GWEALK
B_US_RF_M1   FLTLIWDDLW TLCSFSYHRL RDLLLIVVRI VELLGRR... ....GWEALK
B_US_SF2_K   FLALIWEDLR SLCLFSYRRL RDLLLIAART VEILGHR... ....GWEALK
B_US_WEAU1   FLTLIWVDLR SLCLFLYHRL IDLLLIAKRI VELLGRR... ..........
B_US_WR27_   LLALIWDDLR SLCLFSYHRL RDLLSIVARI VELLGRR... ....GWEILK
B_US_YU2_M   FLAIIWVDLR SLCLFSYHRL RDLLLIVTRI VELLGRR... ....GWGVLK
BF1_BR_93B   FLALVWDDLR NLCLFSYRHL RDFILIAARI VDRGLKR... ....GWEALK
C_BR_92BR0   FLALAWDDLR SLCLFSYHRL RDLILIAARA VELLGRSSLR GIQRGWEILK
C_BW_96BW0   FLALVWDDLR SLCLFSYHRL RDFILIAARA AELLGRSSLR GLQKGWETLK
C_BW_96BW1   FLALAWDDLR SLCLFCYHRL RDFTLVAARA VELLGRSSLK GLQRGWEILK
C_BW_96BW1   FLALAWDDLR SLCLFSYHRL RDSILVAART VELLGRSSLR GLQRGWEALK
C_BW_96BW1   FLALAWDDLR SLCLFSYHRL RDLILVTARV VELLGRSSLR GLQRGWEALK
C_ET_ETH22   FLAIFWDDLR SLCLFSYHRL RDLILIAART VELLGRSSLK GLQRGWETLK
C_IN_93IN1   FLALAWDDLR NLCLFSYHRL RDFISVAARV VELLGRS... ....SWEALK
C_IN_93IN9   FLALAWDDLR NLCLFSYHRL RDFILVAARV VELLGRNSLR GLQRGWEALK
C_IN_93IN9   FLALAWDDLR SLCLFSYHRL RDFILVAVRA VELLGRSSLR GLQRGWEALK
C_IN_94IN1   FLALAWDDLR SLCLFSYHRL RDFILVAARV VELLGHNSLR GLQRGWEALK
C_IN_95IN2   FLALFWDDLR NLCLFSYHRL RDFILVAARV LELLGRRSLR GLQRGWEALK
CRF01_AE_C   FLSLAWDDLR SLCLFSYHRL RDFILIAART VELLGHSSLK GLRRGWEGLK
CRF01_AE_C   FLALAWDDLR SLCLFSYHRL RDFILIAART VELLGHSSLQ GLRRRWEGLK
CRF01_AE_C   FLALVWDDLR SLCLFSYHRL RDFILIAART VELLGHSSLE GLRRGWEGLK
CRF01_AE_T   FLSLAWDDLR SLCLFSYHRL RDFILIATRT VELLGHSSLK GLRRGWEGLK
CRF01_AE_T   FLTLAWDDLR SLCLFSYHRL RDFILIAART VELLGRSSLK GLRRGWEGLK
CRF01_AE_T   FLALAWDDLR SLCLFSYHRL RDFILIAART VELLGHSSLK GLRRGWEGLK
CRF01_AE_T   FLALAWDDLR SLCLFLYHRL RDFILIAART VELLGHSSLK GLRRGWEGLK
CRF01_AE_T   FLALAWDDLR SLCLFSYHRL RDLTLIAART VELLGHSSLK GLRRGWEGLK
```

```
CRF01_AE_T  FLALAWDDLR SLCLFLYHRL RDFILIAART VELLGHSSLK GLRRGWEGLK
CRF02_AG_F  FLALAWDDLR SLCLFSYHRL RDFVLIAVRA VELLGHSSLK GLRLGWEALQ
CRF02_AG_F  FLALAWDDLR SLCLFSYHHL RDFVLIAVRA VELLGHSSLK GLRLGWEALK
CRF02_AG_G  FLALAWDDLR NLCLFSYHRL RDLILIAART VEILGHR... ....VWQILK
CRF02_AG_N  FLALAWDDLR SLCLFSYHRL RDLILIAART VELLGHNCLK GLRLGWGALK
CRF02_AG_S  FLALAWDDLR SLCLFSYHRL RDFVSIVART VELLGHR... ....GWEALK
CRF02_AG_S  FLALAWDDLR SLCLFLYHRL RDFVLIAART VELLGHSSLK GLRLGWEALK
CRF03_AB_R  FLALIWDDLR SLCFFIYHHL RDLLLIAARI VELLGRR... ....GWEALK
CRF03_AB_R  FLALIWDDLR SLCLFIYHHL RDLLLIAART VELLGRR... ....GWEALK
CRF04_cpx_  FLPLIWDDLR NLCLFSYRHL RNLLLIVART VELLGIR... ....GWEALK
CRF04_cpx_  FLPLIWDDLR NLCLFSYHHL RNLLLIAART VELLGRR... ....GWEALK
CRF04_cpx_  FLPLVWDDLR NLCLFSYRQL RNLLLIVAKT VELLGIR... ....GWGTLK
CRF05_DF_B  FSALIWDDLR NLCLFSYHHL RDLTLIVVRI VELLGRR... ....GWEALK
CRF05_DF_B  LSTLIWDDLR NLCLFSYHRL RDLILIAARI VELLGRR... ....GWEALK
CRF06_cpx_  FLALAWDDLR SLCLFSYHRL RDFGLIAART VEILGRR... ....GWEILK
CRF06_cpx_  FLALAWDDLR SLCLFSYHRL RDFVLIAART VGTLGHR... ....GWEILK
CRF06_cpx_  FLALAWDDLR SLCLFSYHRL RDFVLIAART VETLGRR... ....GWEILK
CRF06_cpx_  FLALAWEDLR SLCRFSYHLL RDFVLIVLRT VETLGHR... ....GWEILK
CRF11_cpx_  FLAIAWDDLR NLCLFSYHRL RDFILIVARI VETLGHR... ....GWEILK
CRF11_cpx_  FLALAWDDLR NLCLFLYHQL RDFILIVARI VETLGRR... ....GWESLK
D_CD_84ZR0  FSALIWDDLR NLCLFSYHRL RELILIAARI VELLGRR... ....GWEALK
D_CD_ELI_K  FSALIWDDLR SLCLFSYHRL RDLILIAVRI VELLGRR... ....GWDILK
D_CD_NDK_M  LFALFWDDLR NLCLFSYHRL RDSILIAARI VELLGRR... ....GWEALK
D_UG_94UG1  LSALIWDDLR NLCLFSYHRL RDLILIAARI VELLGRR... ....GWEAIK
F1_BE_VI85  FLALAWDDLR NLCLFSYRHL RDFILIAARI VDRGLRR... ....GWEALK
F1_BR_93BR  FLALAWDDLR NLCLFSYRHL RDFILIAARI VDRGLKR... ....GWEALK
F1_FI_FIN9  FLALVWDDLR NLCLFSYRHL RDFILIAARI VDRGLRR... ....GWEALK
F1_FR_MP41  FLSLVWDDLR NLCLFSYRHL RDFILIAART VDRGLTR... ....GWETLK
F2_CM_MP25  FLALAWDDFR SLCVFSYHCL RNFILIAART VDKGLKR... ....GWEVLK
F2KU_BE_VI  FLALAWDDLR NLCLFSYRHL RDLILIVARI LERGLRG... ....SWEILK
G_BE_DRCBL  FLALAWDDLR SLCLFSYHRL RDFILIAART VELLGRNSLK GLRLGWEALK
G_NG_92NG0  FLALAWDDLR SLCLFSYHRL RDLVLIAART VELLGRSSLK GLRLGWEGLK
G_SE_SE616  FLPLIWDDLR SLCLFSYHRL RDSILIVART VELLGRSSLK GLRLGWEGLK
H_BE_VI991  FLPLVWEDLR NLCLFSYRRL RDLLSIVART VELLGRR... ....GWEALK
H_BE_VI997  FLPIVWDDLR SLCLFSYRLL RDSLLIVIRT VELLGRR... ....GREALK
H_CF_90CF0  FLPVVWDDLR SLSLFSYRLL RDLLLIVVRT VELLGRR... ....GREALK
J_SE_SE702  FLALAWDDLR SLCLFSYHRL RDFVLIAART VGTLGLR... ....GWEILK
J_SE_SE788  FLALAWDDLR NLCLFSYHRL RDFVLIAART VGTLGLR... ....GWEILK
K_CD_EQTB1  FLALAWDDLR NLCLFSYRHL RDLVLIATRI LDRGLKG... ....SWEALK
K_CM_MP535  FLALAWDDLR NLCLFSYRQL RNLILIVTRI LERGLRG... ....GWEALK
N_CM_YBF30  FSALVWEDLR NLLIFLYHRL TDSLLILRRT LELLGQSLSR GLQLLNELRT
O_CM_ANT70  FLPLLYTDLR TIILWTYHLL SNLASGIQKV ISYLRLGLWI LGQKIINVCR
O_CM_MVP51  FLQQLYTDLR TIILWTYHLL SNLISGIRRL IDYLGLGLWI LGQKTIEACR
O_SN_99SE_  FLPLLYTDLR TIILWSYHLL SNLASGIQTV ISHLGLGLWT LGQKIISACR
O_SN_99SE_  FLPLLYTDLR TIILWSYHLL SNLASGIQTV ISHLGLGLWI LGQKIISACR
U_CD___83C  FLALAWEDLR SLCIFSYHRL RDLILIVVKG ...LRR.... ....GWEALK
```

```
                901                                                  950
00BW0762_1  YLGILVQYWG LELKKSAISL FDTIAIAVAE GTDRIIEAIQ RICRAICNIP
00BW0768_2  YLGNLVLYWG LELKKSAISL LDSIAIAVAE GTDRILEAVQ RIWGAIRNIP
00BW0874_2  YLGSLVQYWG LELKKSAISL LDTIAIAVAE GTDRIIELIQ RICRAIYNIP
00BW1471_2  YLGSLGQYWG QELKKSAINL FDTIAIAVAE GTDRIIEAVQ RAVRAILHIP
00BW1616_2  YLGSLVQYWG LELKKSAVSL LDTIAIAVAE GTDRILEVTQ RICRVIRNIP
00BW1686_8  YLGSLIQYWG LELKKSAISL LDTIAIAVAG GTDRFIELIQ RIYRAIRNVP
00BW1759_3  YLGSLGQYWG LELKKSAISL LDTIAIAVAE GTDRIIELIQ TICRAIRNIP
00BW1773_2  YLGNLVQDWG LELKKSAISL FDAIAIAVAE GTDRIIELIQ RTGRAICNIP
00BW1783_5  YLGTLVQYWV LELKKSAISL LDATAITVAG GTDRIIELIQ RIGRAILSIP
00BW1795_6  YLGSLVQYWG LELKKSAISL LDTVAIAVAE GTDRIIELIQ RGYRAICNIP
00BW1811_3  YLGSLVQYWG LELKKSAISL LDTIAIAVGE GTDRIIEIIQ RICRAIRNTP
00BW1859_5  YLGSLVQYWG LELKKSAISL LDTIAIAVAE GTDRIIDLIQ RICRAILRIP
00BW1880_2  YLGSLIQYWG LELKKSAISL LDTIAIAVAE GTDRIIEGIQ RICRIIRNIP
00BW1921_1  YLGSLIQYWG LELKKSAISL LDTIAIATAE GTDRIIEVIQ RICRVIRNIP
00BW2036_1  YLGSLVQYWG LELKKSAISL LDTIAIAVAE GTDRIIELVQ RIGRGIYNIP
00BW2063_6  YLGSLVQYWG LELKKSAISL LNTTAIAVAE GTDRVIELLQ RIGRAICNIP
00BW2087_2  YLGSLVQYWG LELRKSASSL LDTIAIAVAE GTDRIIEVIQ IICRAILHIP
00BW2127_2  YLGNLVLYWG LELKKSAISL FDTIAVAVAE GTDRILEVIQ RICRAIRNIP
00BW2128_3  YLGSLVQYWG LELKKSAVSL LNTIAIVVAE GTDRILELIQ RLRRAFLNIP
00BW2276_7  YLGNLAQYWG LELKKSAISL INTIAIAVGE RTDRIIELIQ TLCRAIHNIP
00BW3819_3  YLGNLVQYWG LELKRSAISL LDTIAIAVAE GTDRIIEFLQ RIFRAIRNIP
00BW3842_8  YLGNLVQYWG LELKKSAISL LDAIAIAVGE GTDRILELLQ RIGRGICNIP
00BW3871_3  YLGSLIQYWG LELKKSAINL LDTTAIAVAE GTDRFIELIQ RICRAVRNIP
00BW3876_9  YLKNLGLYWG LELKKSAISL LNTIAIAVAE GTDRVIEFVL RICRAIRHIP
00BW3886_8  YLGSLVQYWG LELKKSATSL LDTIAIAVAE GTDRIIETVL RICRAILHIP
00BW3891_6  YLGSLVQYWG LELKKSAISL LDTIAIVVAE GTDRIIELVL GICRAIRNVP
00BW3970_2  YLASLVQYWG LELKKGAISL LDSIAIAVAE GTDRIIAFIQ RLFRAICNLP
00BW5031_1  YLGSLVQYWG LELKKSAISL LDTIAIAVAE GTDRIIEVVQ RLYRAILNIP
 96BW01B21  YLGNLLLYWG LEPKKSAINL LDTTAIAVAE GTDRILELVQ GICRAIRNIP
  96BW0407  FLGSLVQYWG LELKKSAISL LDTTAIAVAE GTDRIIEIAQ RICRAICNVP
  96BW0502  YLGSLVQYWG LELKKSAISL LDTIAIAVAE GTDRIIEFIQ RICRAIRNIP
 96BW06_J4  YLGSLIQYWG LELKRSTISL LDTVPIAVPE GTDRIIELIQ RIWRAICNIP
 96BW11_06  YLGSLVQYWG LELKKSAISL LDTTAIAVAE GTDRIIEVLQ RIGRAIRNTP
  96BW1210  YLGSLVQYWG LELKKSAISL LDTIAIAVAE GTDRIIELTQ RVFRAIRNIP
 96BW15B03  YLGSLVQYWG LELKKSATSL LDSIAIAVAE GTDRIIEVIQ RIYRAFCNIP
 96BW16_26  YLGSLVQYWG LELKKSAINL LDTIAIAVAE GTDRIIDFIL RICRAIRNIP
 96BW17A09  YLGSLGQYWG QELKKSAINL LDTIAIAVAE GTDRIIEVLQ GAIRAILNIP
 96BWMO1_5  YLGSLVQYWG LELKKSAISL LDTTAIAVAE GTDRIIEVLQ RVGRAIRNTP
 96BWMO3_2  YLGSLVRYWG LELKKSAISL LDTIAVAVAE GTDRIIEVIQ GICRGIRNIP
98BWMC12_2  YLGSLVQYWG LELKKRAISL LDTTAIAVAE GTDRIIEIVL RICRAICNVR
98BWMC13_4  YLGSLVQYWG LELKKSAISL LDTTAIAVAE GTDRIIELLQ RIGRAIRNTP
98BWMC14_a  YLGNLIQYWG LELKKSAINL LDTLAIAVAE GTDRIIELIQ RVCRAILNIP
98BWMO14_1  YLGNLVQYWG LKLKKSAISL FDTIAIAVAE GTDRIIELIQ IICRAIRNIP
98BWMO18_d  YLGSLVQYWG LELKKSAISL LDTIAIAVAE GTDRIIELVQ RICRGVLNIP
98BWMO36_a  YLGSLVQYWG LELKKSAISL LDTIAIATAE GTDRIIELIQ RICRAIYNIP
98BWMO37_d  YLGNLVQYWG LELKKSAISL LDTIAIAVAE GTDRIIEFIQ RICRAIRNLP
99BW3932_1  YLGSLVQYWG LELKKSAISL LDATAVAVAE GTDRILEIIQ RIFRAICNIP
99BW4642_4  YLGSLVQYWC LELKKSATSL IDAIAIAVAE GTDRIIDLIQ RICRAIRNIP
99BW4745_8  YLGSLVQYWG LELKKSAISL FDTIAIAVAE GTDRIIELVL RICGAIRNIP
99BW4754_7  YLGSIVQYWG LKLKKSAISL LDTTAIAVAE GTDRIIELLR RFCRAIYSIP
99BWMC16_8  YLGSLGQYWG LELKKSAIGL LDTIAIAVAE GTDRIIELIQ RTFRAICNIP
A2_CD_97CD  HLWNLLVYWG QELKTSAIRL LDTIAVAVAE WTDRVIEIGQ RACRAIRNIP
A2_CY_94CY  NLWNLLLYWG RELKNSAISL FDTIAVAVAE WTDRVIEIGQ RAFRAILNIP
A2D___97KR  YLWNLLLYWG RELKNSAISL FNATAIAVAE WTDRVIEIVQ RACRAIINIP
A2G_CD_97C  YLWNLLLYWG QELKNSASNL LDTVAIAVAN WTDRVIEAAQ GACRAIRNVP
A_BY_97BL0  YXXNLXGYXG QELKSSAINL IDTIAIAVAX XTDXVIEIGQ RFCRAIRNIP
A_KE_Q23_A  YLWNLLSYWG RELKISAINL VDTIAIAVAG WTDRVIEIAQ RIGRAILHIP
A_SE_SE659  YLGNLLLYWG RELKISAINL LDTTAIAVAG WTDRVIEIVQ GIGRAFLHIP
A_SE_SE725  YLGNLLLYWG QELKLSAISL FDTPAIAVAG WTDRGIELIQ RIGRAILNIP
A_SE_SE753  YLWNLLLYWG RELKSSAINL VDTIAIAVAG WTDRIIEIGL RIGRAFLHIP
```

```
A_SE_SE853   YLWNLLVYWI RELKISAISL LDTIAIAVAG WTDRVIELGQ RLCRAILHIP
A_SE_SE889   YLKNLLSYWG RELKLSAINL LDTIAIVIAG WTDRVIEIGQ GFCRAIFHP.
A_SE_UGSE8   YLGNLLLYWI RELKISAISL FDTIAIAVAG WTDRVIEIGQ RIGRAILHIP
A_UG_92UG0   YLGNLLLYWG RELKISAINL LDTIAIAVAG WTDRVIETVQ RLGRAILNIP
A_UG_U455_   YLWNLLLYWG RELKISAITL LDAVAVAVAG WIDRVIEIGQ TIGRAILNIP
AC_IN_2130   YLWNLLVYWG RELKISAIKL VDTIAIVVAG WTDRIIEIGQ GIGRAILHIP
AC_RW_92RW   YLGNLVQYWG LELKRSAINL LDTTAIVVAE GTDRIIELIQ RISRAIYNIP
AC_SE_SE94   YLWNLLLYWG RELRISAINL LDTIAIATAS WTDRVIELGQ RICRAILNIP
ACD_SE_SE8   YLWNLLQYWI QELKNSAINL FNTIAIAVAE GTDRVIEIGQ RIGRAILNTP
ACG_BE_VI1   LLGNILLYWS QELKNSAINL LDTIAIAVAN WTDRVIEIGQ RAGRAFLNIP
AD_SE_SE69   YLWNLLQYWI QELKNSAISL VDTTAIAVAE GTDRVIETVQ RAFRAVLRIP
AD_SE_SE71   YLWNLLQYWI QELKISAISL VDSIAIVVAG WTDRVIEIGQ GIGRAILHIP
ADHK_NO_97   YLGNLLLYWG QELKNSAINL LNTTAIAVAE GTDRIIEIVQ RTGRAVLHIP
ADK_CD_MAL   YLWNLLQYWG QELKNSAISL LNTTAIAVAE CTDRVIEIGQ RFGRAILHIP
AG_BE_VI11   YLWNLLVYWG QELKNSAINL LDTVAIAVAN WTDRVIEIGQ RAGRAILNIP
AG_NG_92NG   YLWNLLLYWG RELKNSAINL IDTIAIAVAN WTDRVIEVAQ GACRAILNIP
AGHU_GA_VI   YLWNLLLYWG QELKSSAISL LDAVAIAVAN WTDRVIEVVQ RVGRAILNIP
AGU_CD_Z32   YLGNLVIYWG QELKNSAINL LDTVAIAVAD WTDRVIEVVQ RAGRAFLNIP
AJ_BW_BW21   YLGNLALYWG RELKNSAISL LDTIAITVAE ATDRIIEIAQ RAFRAILHIP
B_AU_VH_AF   YWWNLLQYWS QELQNSAISL LNATAIAVAE GTDRVIEVVQ RACRAILHIP
B_CN_RL42_   YWWNLLQYWI QELKNSAIGL LNATAIAVAE GTDRVIEVVQ RAYRAILHIP
B_DE_D31_U   YWWNLLQYWS QELKNSAVSL LNATAIAVAE GTDRVIEVVQ RAWRAILHIP
B_DE_HAN_U   YWWNLLQYWS QELKNSAVSL FNTIAIAVAE GTDRVIEVVQ RACRAILHIP
B_FR_HXB2_   YWWNLLQYWS QELKNSAVSL LNATAIAVAE GTDRVIEVVQ GACRAIRHIP
B_GA_OYI___  YWWNLLQYWS QELKNSVISL LNATAIAVAE GTDRVIEIVQ RAYRAFLNIP
B_GB_CAM1_   YWWNLLQYWS QELRNSAVSL FDTIAIAVAE GTDRVIEVVQ RACRAILHIP
B_GB_GB8_C   YWWNLLQYWI QELKNSAISL LNTTAIAVAE GTDRVIEVVQ RAYRAILHIP
B_GB_MANC_   YWWNLLQYWS QVLKNSAVSL LNVTAIAVAE GTDRIIEVVQ RVGRAILHIP
B_KR_WK_AF   YLWNLLQYWS QELKNSAVSL LNATAVAVAE GTDRIIEILQ RAYRAILNIP
B_NL_3202A   YWWNLLQYWS. QELKNSAVSL LNATAIAVAE GTDRVIEVVQ RACRAVLHIP
B_TW_TWCYS   YLWNLLQYWI QELKNSAVSL FNAIAIAVAE GTDRVIEVVQ RVFRAILHIP
B_US_BC_L0   YWWSLLQYWS QELKNSAVNL LNVTAIAVAE GTDRVIEVVQ RTYRAILHIP
B_US_DH123   YLWNLLQYWS QELKNSAVSL LNATAIAVGE GTDRIIEILQ RAGRAILNIP
B_US_JRCSF   YWWNLLQYWS QELKNSAVSL LNATAIAVAE GTDRIIEVVQ RVYRAILHIP
B_US_MNCG_   YWWNLLQYWS QELKSSAVSL LNATAIAVAE GTDRVIEVLQ RAGRAILHIP
B_US_P896_   YWWNLLQYWS QELKNSAVSL LNATAIAVAE GTDRVIKIVQ RACRAIRNIP
B_US_RF_M1   YWWNLLQYWS QELKNSAVSL LNTTAIAVAE GTDRIIEVAQ RILRAFLHIP
B_US_SF2_K   YWWSLLQYWI QELKNSAVSW LNATAIAVTE GTDRVIEVAQ RAYRAILHIH
B_US_WEAU1   .......... .......... .......... .......... ..........
B_US_WR27_   YWGNLLQYWG QELRNSAISL LNATAIAVAE GTDRVIEVGQ RIFRAILHIP
B_US_YU2_M   YWWNLLQYWI QELKNSAVSL LNATAIAVAE GTDRVIEILQ RAFRAVLHIP
BF1_BR_93B   LLGNLALYWS QELKNSAISL LNTTAIVVAE GTDRVIEALQ RAGRAVLNVP
C_BR_92BR0   YLGGLVQYWS LELKKSAISL FDTIAIAVAE GTDRIIEVIQ GIWRAICNIP
C_BW_96BW0   FLGSLVQYWG LELKKSAISL LDTTAIAVAE GTDRIIEIAQ RICRAICNIP
C_BW_96BW1   YLGSLVQYWG LELKMSTISL LDTTAIAIAE GTDRIIELIQ RIGRAIRNTP
C_BW_96BW1   YLGSLVQYWG LELKKSAISL LDTIAIAVAE GTDRIIELTQ RVFRAIRNIP
C_BW_96BW1   YLGSLVQYWG LELKKSATSL LDSIAIAVAE GTDRIIEVIQ RIYRAFCNIP
C_ET_ETH22   YLGSLVQYWG LELKKSAINL LNTTAIVVGE GTDRFIELIQ RIWRAFCNIP
C_IN_93IN1   YLGSLVQYWG LELKKSAISL FDSIAIVVAE GTDRIIELVQ GFCRAIRNIP
C_IN_93IN9   YLGSLVQYWG LELKKSAISL LDIIAIAVAE GTDRIIELIQ RTCRAIRNIP
C_IN_93IN9   YLGSLVQYWG IELKRSAISL LDFTAIAVAE GTDRIIELVL RICRAIRNIP
C_IN_94IN1   YLGSLVQYWG LELKKSAIRL LDIIAIAVAE GTDRIIEIIQ GTCRAIRNIP
C_IN_95IN2   YLGSLVQYWG LELKKSAINL LDRIAIAVAE GTDRILELVQ RICRAIRNIP
CRF01_AE_C   YLGNLLSYWG QELKTSAITL FDAIAVAVAG WTDRVIEVVQ RAWRALIHIP
CRF01_AE_C   YLGNLLSYWV QELRISAITL LDATAITVAG WTDRVIEIVQ RAWRAILHIP
CRF01_AE_C   YLGSLLSYWG QELKTSAITL LDATAITVAG WTDRAIEIAQ RACRAILHIP
CRF01_AE_T   YLGNLLLYWG QELKISAISL LNTTAIAVAG WTDRVIEVAQ GAWRAILHIP
CRF01_AE_T   YLGNLLLYWG QELKISAISL LDATAIAVAG WTDRVIEVAQ GAWRAILHIP
CRF01_AE_T   YLGNLLLYWG QELKISAISL FDALAVVVAG WTDRVIEVAQ GAWRAILHIP
CRF01_AE_T   YLGNLLLYWG QELKISAISL LDAIAIAVAG WTDRVIKVAQ RAWRAILHIP
CRF01_AE_T   YLGNLLLYWG QELKISAISL LDATAIAAAG WTDRVIEVAQ GAWRAILHIP
```

```
CRF01_AE_T  YLWNLLVYWG QELKISVISL LNATAIVVAG WTDRVIEVAQ GAWRAILHIP
CRF02_AG_F  YLGNLLLYWG QELKNSAINL LDTIALAVAN WTDRVIEIGQ RVGRAILNIP
CRF02_AG_F  YLGNLLTYWG QELKNSAINL LDTIAIAVAN WTDRVIEIGQ RVGRAIRNIP
CRF02_AG_G  YLGNLAQYWG LELKNSAISL LNTTAIVVAE QTDRLLEFLQ RAGRAILHIP
CRF02_AG_N  YLWNLISYWV QELKNSAINL LNTIAIVVAN WTDRAIEIGQ RVGRAIRNIP
CRF02_AG_S  YLWNLLSYWG QELKNSAISL LDTIAIVVAN WTDRVIELVQ RAGRAILNIP
CRF02_AG_S  HLWNLLSYWG QELKNSAINL LDTTAVAVAN WTDRVIEIVQ RTGRAICNIP
CRF03_AB_R  YWWNLLQYWI QELKSSAINL IGTIAIAVAG WTDRVIEIGQ RFCRAMRNIP
CRF03_AB_R  YWWNLLQYWI QELKSSAINL INTIAIAVAG WTDRVIEIGQ RFCRAIRNIP
CRF04_cpx_  YLWNFLLYWG QELKNSAINL FNTTAIAVAE GTDRIIEAVQ RACRAICNIP
CRF04_cpx_  YLWNLLLYWG QELRNSAINL LDTTAIAVAE GTDRILEAVQ RACRAIRNIP
CRF04_cpx_  YLWNLLLYWG QEIRSSAISL LDTTAVAVAE GTDRIIEAVQ RICRAILNIP
CRF05_DF_B  YLWSLPQYWS RELKNSAISL LNTTAVVVAE GTDRVIEALQ RAGRAILNIP
CRF05_DF_B  YLWSLLQYWS QELKNSAISL LNTTAVVVAE GTDRILEALQ RAGRAVLNIP
CRF06_cpx_  YLGNLICYWG QELQNSAISL FDAAAIAVAN WTDRVAEVVQ RIFRAFLNVP
CRF06_cpx_  YLGSLVWYWG QELKNSAINL LDTTAIAVAN GTDRVIEIVQ RAFRAVLNIP
CRF06_cpx_  YLWNLVCYWG QELKNSAISL IDTTAIAVAN WTDRVIEVVQ RAFRAVLNIP
CRF06_cpx_  YLGNLVCYWG QELKNSAISL LDTTAIAVAN WTDRVIEIVQ RVFRAFLNVP
CRF11_cpx_  YLGNLTQYWG QELKNSAINL LNTTAIAVAE GTDRIIEIVQ RVLRGILHIP
CRF11_cpx_  YLGNLAQYWG QELKSSAISL LNATAIAVAE GTDRIIEVAH RALRAILNIP
D_CD_84ZR0  YLWNLLQYWS RELKNSAISL VDATAIAVAE GTDRIIDIVR RACKAVLHIP
D_CD_ELI_K  YLWNLLQYWS QELRNSASSL FDAIAIAVAE GTDRVIEIIQ RACRAVLNIP
D_CD_NDK_M  YLWNLLQYWS QELRNSASSL LDTIAIAVAE RTDRVIEVVQ RACRAILNVP
D_UG_94UG1  YLWNLLQYWI QELKNSAVSL FNTIAIAVAE GTDRAIELVQ RAVRAILNIP
F1_BE_VI85  YLGNLTRYWS QELKNSAISL FNTTAIVVAE GTDRIIEVLQ RAGRAVLNIP
F1_BR_93BR  YLGNLTQYWG QELKNSAISL LNATAIAVAE WTDRVIEALQ RAGRAILNIP
F1_FI_FIN9  YLGNIIQYWS QELKNSAISL FNTTAIVVAE GTDRVIEALQ RAVRAVLNIP
F1_FR_MP41  YLWNLAQYWS QELKNSAISL LNTTAIVVAE GTDRVIEVLQ RAGRAVLNVP
F2_CM_MP25  YLWNLAQYWG QELKNSAISL LDRTAIAVAE GTDRIIEILQ RAGRAVLNIP
F2KU_BE_VI  YLWSLVQYWG QELKNSAINL LNTTAIAVAE GTDRIIEVFQ RAGRAVLNIP
G_BE_DRCBL  YLWNLLLYWA RELKNSAINL LDTIAIAVAN WTDRVIEVAQ RAGRAVLNIP
G_NG_92NG0  YLWNLLLYWG RELKNSAINL LDTIAIATAN GTDRVIEVAQ RAYRAILNVP
G_SE_SE616  YLWNLLLYWG RELKNSAISL LDTVAIAVAN WTDRVIEVAQ RACRAILNIP
H_BE_VI991  LLGNLLLYWG QELKNSAISL LNTTAIAVAE GTDRIIELVQ RAWRAILHIP
H_BE_VI997  YLWNLLQYWG QELKNSAINL LNTTAIVVAE GTDRIIEIVQ RAWRAVLHIP
H_CF_90CF0  YLWNLLQYWG QELKNSAIDL LNTTAIAVAE GTDGIIVIVQ RAWRAILHIP
J_SE_SE702  YLVNLVWYWG QELKNSAISL LNTTAIAVAE GTDRIIEIAQ RAFRAILHIP
J_SE_SE788  YLVNLVWYWG QELKNSAISL LNTTAIAVAE GTDRIIEIAQ RAFRAILHIP
K_CD_EQTB1  YLWNLILYWG QEIKNSAINL LNTTAIAVAE GTDRIIEIVY RAFRALLHIP
K_CM_MP535  YLWNLVQYWS QELKNSAISL LNTTAIAVAG GTDRIIEIGQ RAFRALLHIP
N_CM_YBF30  HLWGILAYWG KELRDSAISL LNTTAIVVAE GTDRIIELAQ RIGRGILHIP
O_CM_ANT70  ICAAVTQYWL QELQNSATSL LDTLAVAVAN WTDGIIAGIQ RIGTGIRNIP
O_CM_MVP51  LCGAVMQYWL QELKNSATNL LDTIAVSVAN WTDGIILGLQ RIGQGFLHIP
O_SN_99SE_  LCIAVIQYWL QELQNSATSL LDTIAVAVAN WTVTIILGIQ RIGRGILNIP
O_SN_99SE_  ICIAVIQYWL QELQNSATSL LDTLAVAVAN WTDGIILGLQ RIGRGILNIP
U_CD___83C  YLGNLVLYWG QELKNSAISL LNATAIVVAE GTDRIIEVGQ RICRAILNIP
```

	951	962
00BW0762_1	RRIRQGFEAA	LQ
00BW0768_2	RRIRQGFEAA	LQ
00BW0874_2	RRIRQGFEAA	LQ
00BW1471_2	RRIEQTFEPP	LL
00BW1616_2	RRIRQGVEAA	LQ
00BW1686_8	RRIRQGFETA	LL
00BW1759_3	RRIRQGFEAA	LQ
00BW1773_2	RRIRQGFEAA	LQ
00BW1783_5	RRIRQGFEAA	LQ
00BW1795_6	TRIRQGFEAA	LQ
00BW1811_3	RRIRQGFEAS	LL
00BW1859_5	RRIRQGFEAA	LL
00BW1880_2	TRIRQGFEAA	LQ
00BW1921_1	TRIRQGFEAA	LQ
00BW2036_1	RRIRQGFEAA	LQ
00BW2063_6	RRVRQGFETA	LL
00BW2087_2	RRIRQGFEVA	LL
00BW2127_2	TRIRQGFEAA	LL
00BW2128_3	SSIRQGFEAA	LQ
00BW2276_7	RRIRQGFEAA	LL
00BW3819_3	TRIRQGFEAA	LL
00BW3842_8	RRIRQGFEAA	LQ
00BW3871_3	RRLRQGFEAA	LL
00BW3876_9	RRIRQGFEAA	LL
00BW3886_8	RRIRQGFEAA	LL
00BW3891_6	TRIRQGFEAA	LQ
00BW3970_2	RRIRQGFEAS	LL
00BW5031_1	RRIRQGFEAA	LQ
96BW01B21	RRIRQGFEAA	LQ
96BW0407	TRIRQGFEAA	LQ
96BW0502	RRIRQGFEAA	LQ
96BW06_J4	RRIRQGFEAA	LL
96BW11_06	RRIRQGFETA	LL
96BW1210	RRIRQGFEAA	LQ
96BW15B03	RRVRQGFEAA	LQ
96BW16_26	RRLRQGFEAA	LQ
96BW17A09	TRIRQGLEAA	LQ
96BWMO1_5	RRIRQGFEAA	LL
96BWMO3_2	RRIRQGFEAA	LL
98BWMC12_2	.....GFEAA	LQ
98BWMC13_4	RRIRQGFETA	LL
98BWMC14_a	RRVRQGFEAA	LQ
98BWMO14_1	TRIRQGLEAA	LL
98BWMO18_d	RRIRQGFEAA	LQ
98BWMO36_a	TRIRQGFEAA	LL
98BWMO37_d	RRIRQGFEAA	LL
99BW3932_1	RRIRQGFETA	LL
99BW4642_4	RRIRQGFEAA	LQ
99BW4745_8	TRIRQGFEAA	LQ
99BW4754_7	RRIRQGFEAA	LQ
99BWMC16_8	RRIRQGFETA	LL
A2_CD_97CD	RRIRQGLERA	LL
A2_CY_94CY	RRIRQGLERA	LL
A2D___97KR	RRIRQGLERA	LL
A2G_CD_97C	RRIRQGLERA	LL
A_BY_97BL0	RRIRXGAEKA	LQ
A_KE_Q23_A	VRIRQGLERA	LL
A_SE_SE659	RRIRQGFERA	LL
A_SE_SE725	RRIRQGFEEA	LL
A_SE_SE753	RRIRQGFERA	LL

```
A_SE_SE853   VRIRQGFERA LL
A_SE_SE889   RRSKQGLKRA LQ
A_SE_UGSE8   RRIRQGFER. ..
A_UG_92UG0   RRIRQGFERA LL
A_UG_U455_   RRIRQGLERA LL
AC_IN_2130   RRIRQGLERA LL
AC_RW_92RW   SRIRQGFEAA LQ
AC_SE_SE94   RRIRQGFERA LL
ACD_SE_SE8   RRIRQGLERA LL
ACG_BE_VI1   RRIRQGFERA LL
AD_SE_SE69   ARIRQGLERV LL
AD_SE_SE71   RRIRQGLERA LL
ADHK_NO_97   RRIRQGFERX LL
ADK_CD_MAL   RRIRQGFERA LL
AG_BE_VI11   RRIRQGLERA LL
AG_NG_92NG   RRIRQGLERA LL
AGHU_GA_VI   RRIRQGLERA LI
AGU_CD_Z32   RRIRQGLERA LL
AJ_BW_BW21   VRIRQGFERA LL
B_AU_VH_AF   RRIRQGLERL LL
B_CN_RL42_   TRIRQGLERA LL
B_DE_D31_U   VRIRQGLERA LL
B_DE_HAN_U   RRVRQGLERA LL
B_FR_HXB2_   RRIRQGLERI LL
B_GA_OYI__   RRIRQGLERA LL
B_GB_CAM1_   RRIRQGLERL LL
B_GB_GB8_C   TRIRQGLERA LQ
B_GB_MANC_   VRIRQGLERA LL
B_KR_WK_AF   RRIRQGLERA LL
B_NL_3202A   VRIRQGLERA LL
B_TW_TWCYS   TRIRQGLERA LL
B_US_BC_L0   RRIRQGLERL LL
B_US_DH123   TRIRQGLERA LL
B_US_JRCSF   TRIRQGLERA LL
B_US_MNCG_   TRIRQGLERA LL
B_US_P896_   TRIRQGLERA LL
B_US_RF_M1   RRIRQGLERA LL
B_US_SF2_K   RRIRQGLERL LL
B_US_WEAU1   .......... ..
B_US_WR27_   RRIRQGLERV LL
B_US_YU2_M   VRIRQGLERA LL
BF1_BR_93B   RRIRQGLERA LL
C_BR_92BR0   RRIRQGFEAA LQ
C_BW_96BW0   TRIRQGFEAA LQ
C_BW_96BW1   RRIRQGFETA LL
C_BW_96BW1   RRIRQGFEAA LQ
C_BW_96BW1   RRVRQGFEAA LQ
C_ET_ETH22   RRIRQGLEAA LQ
C_IN_93IN1   TRIRQGFEAA LQ
C_IN_93IN9   RRIRQGFEAV LQ
C_IN_93IN9   TRIRQGFEIA LQ
C_IN_94IN1   RRIRQGLEAA LQ
C_IN_95IN2   RRIRQGFEAA LQ
CRF01_AE_C   RRIRQGLERA LL
CRF01_AE_C   RRIRQGLERA LL
CRF01_AE_C   RRIRQGLERA LL
CRF01_AE_T   RRIRQGLERT LL
CRF01_AE_T   RRIRQGLERA LL
CRF01_AE_T   RRIRQGLERA LL
CRF01_AE_T   RRIRQGLERA LL
CRF01_AE_T   RRIRQGLERT LL
```

```
CRF01_AE_T  RRIRQGLERA LL
CRF02_AG_F  RRIRQGLERA LL
CRF02_AG_F  VRIRQGLERA LL
CRF02_AG_G  RRIRQGFERA LL
CRF02_AG_N  RRIRQGFERA LL
CRF02_AG_S  RRIRQGFERA LL
CRF02_AG_S  RRIRQGLERA LQ
CRF03_AB_R  RRIRQGAEKA LQ
CRF03_AB_R  RRIRQGAEKA LQ
CRF04_cpx_  RRIRQGLERA LL
CRF04_cpx_  RRIRQGFEKA LL
CRF04_cpx_  RRIRQGLERA LL
CRF05_DF_B  RRIRQGLERA LL
CRF05_DF_B  RRIRQGLERA LL
CRF06_cpx_  RRIRQGFERA LL
CRF06_cpx_  TRIRQGFERA LL
CRF06_cpx_  RRIRQGAERA LI
CRF06_cpx_  RRIRQGFERA LL
CRF11_cpx_  RRIRQGLERA LL
CRF11_cpx_  RRIRQGFERA LL
D_CD_84ZR0  TRIRQGLERA LL
D_CD_ELI_K  RRIRQGLERS LL
D_CD_NDK_M  RRIRQGLERL LL
D_UG_94UG1  VRIRQGLERA LL
F1_BE_VI85  RRIRQGAERA LL
F1_BR_93BR  RRIRQGLERA LL
F1_FI_FIN9  RRIRQRVERA LI
F1_FR_MP41  RRIRQGLERS LL
F2_CM_MP25  RRIRQGLERA LL
F2KU_BE_VI  RRIRQGFERA LL
G_BE_DRCBL  RRIRQGLERA LL
G_NG_92NG0  TRIRQGLERA LL
G_SE_SE616  TRIRQGLERA LL
H_BE_VI991  RRIRQGFERA LL
H_BE_VI997  RRIRQGLERI LL
H_CF_90CF0  RRIRQGFERS LL
J_SE_SE702  RRIRQGLERA LL
J_SE_SE788  RRIRQGLERA LL
K_CD_EQTB1  RRIRQGFERL LL
K_CM_MP535  RRIRQGLERA LL
N_CM_YBF30  RRIRQGLERA LI
O_CM_ANT70  RRIRQGLERS LL
O_CM_MVP51  RRIRQGAERI LV
O_SN_99SE_  RRIRQGLERS LL
O_SN_99SE_  RRIRQGLERA LL
U_CD___83C  RRIRQGFERA LL
```

Table 13. HIV Nef Sequence Alignment

GCG Multiple Sequence File.
Written by Omiga 1.1

Name:	00BW0762_1	Len:	232	Check:	3461	Weight:	1.00
Name:	00BW0768_2	Len:	232	Check:	5650	Weight:	1.00
Name:	00BW0874_2	Len:	232	Check:	3483	Weight:	1.00
Name:	00BW1471_2	Len:	232	Check:	9491	Weight:	1.00
Name:	00BW1616_2	Len:	232	Check:	1504	Weight:	1.00
Name:	00BW1686_8	Len:	232	Check:	1380	Weight:	1.00
Name:	00BW1759_3	Len:	232	Check:	5319	Weight:	1.00
Name:	00BW1773_2	Len:	232	Check:	156	Weight:	1.00
Name:	00BW1783_5	Len:	232	Check:	8063	Weight:	1.00
Name:	00BW1795_6	Len:	232	Check:	3123	Weight:	1.00
Name:	00BW1811_3	Len:	232	Check:	4460	Weight:	1.00
Name:	00BW1859_5	Len:	232	Check:	9116	Weight:	1.00
Name:	00BW1880_2	Len:	232	Check:	4302	Weight:	1.00
Name:	00BW1921_1	Len:	232	Check:	2737	Weight:	1.00
Name:	00BW2036_1	Len:	232	Check:	4558	Weight:	1.00
Name:	00BW2063_6	Len:	232	Check:	1020	Weight:	1.00
Name:	00BW2087_2	Len:	232	Check:	7532	Weight:	1.00
Name:	00BW2127_2	Len:	232	Check:	3425	Weight:	1.00
Name:	00BW2128_3	Len:	232	Check:	5136	Weight:	1.00
Name:	00BW2276_7	Len:	232	Check:	3623	Weight:	1.00
Name:	00BW3819_3	Len:	232	Check:	993	Weight:	1.00
Name:	00BW3842_8	Len:	232	Check:	6030	Weight:	1.00
Name:	00BW3871_3	Len:	232	Check:	3547	Weight:	1.00
Name:	00BW3876_9	Len:	232	Check:	1951	Weight:	1.00
Name:	00BW3886_8	Len:	232	Check:	3786	Weight:	1.00
Name:	00BW3891_6	Len:	232	Check:	3655	Weight:	1.00
Name:	00BW3970_2	Len:	232	Check:	8913	Weight:	1.00
Name:	00BW5031_1	Len:	232	Check:	2223	Weight:	1.00
Name:	96BW01B21	Len:	232	Check:	2176	Weight:	1.00
Name:	96BW0407	Len:	232	Check:	5261	Weight:	1.00
Name:	96BW0502	Len:	232	Check:	333	Weight:	1.00
Name:	96BW06_J4	Len:	232	Check:	5784	Weight:	1.00
Name:	96BW11_06	Len:	232	Check:	4950	Weight:	1.00
Name:	96BW1210	Len:	232	Check:	6118	Weight:	1.00
Name:	96BW15B03	Len:	232	Check:	5089	Weight:	1.00
Name:	96BW16_26	Len:	232	Check:	3957	Weight:	1.00
Name:	96BW17A09	Len:	232	Check:	1945	Weight:	1.00
Name:	96BWMO1_5	Len:	232	Check:	5827	Weight:	1.00
Name:	96BWMO3_2	Len:	232	Check:	2303	Weight:	1.00
Name:	98BWMC12_2	Len:	232	Check:	2423	Weight:	1.00
Name:	98BWMC13_4	Len:	232	Check:	4043	Weight:	1.00
Name:	98BWMC14_a	Len:	232	Check:	3568	Weight:	1.00
Name:	98BWMO14_1	Len:	232	Check:	4909	Weight:	1.00
Name:	98BWMO18_d	Len:	232	Check:	3505	Weight:	1.00
Name:	98BWMO36_a	Len:	232	Check:	6393	Weight:	1.00
Name:	98BWMO37_d	Len:	232	Check:	1912	Weight:	1.00
Name:	99BW3932_1	Len:	232	Check:	19	Weight:	1.00
Name:	99BW4642_4	Len:	232	Check:	6848	Weight:	1.00
Name:	99BW4745_8	Len:	232	Check:	938	Weight:	1.00
Name:	99BW4754_7	Len:	232	Check:	1379	Weight:	1.00
Name:	99BWMC16_8	Len:	232	Check:	4222	Weight:	1.00
Name:	A2_CD_97CD	Len:	232	Check:	2359	Weight:	1.00
Name:	A2_CY_94CY	Len:	232	Check:	5163	Weight:	1.00
Name:	A2D___97KR	Len:	232	Check:	9468	Weight:	1.00
Name:	A2G_CD_97C	Len:	232	Check:	4189	Weight:	1.00
Name:	A_BY_97BL0	Len:	232	Check:	2590	Weight:	1.00
Name:	A_KE_Q23	Len:	232	Check:	2652	Weight:	1.00

Name: A_SE_SE659 Len: 232 Check: 9245 Weight: 1.00
Name: A_SE_SE725 Len: 232 Check: 985 Weight: 1.00
Name: A_SE_SE753 Len: 232 Check: 1638 Weight: 1.00
Name: A_SE_SE853 Len: 232 Check: 2503 Weight: 1.00
Name: A_SE_SE889 Len: 232 Check: 2327 Weight: 1.00
Name: A_SE_UGSE8 Len: 232 Check: 9538 Weight: 1.00
Name: A_UG_92UG0 Len: 232 Check: 2621 Weight: 1.00
Name: A_UG_U455 Len: 232 Check: 2084 Weight: 1.00
Name: AC_IN_2130 Len: 232 Check: 2406 Weight: 1.00
Name: AC_RW_92RW Len: 232 Check: 3441 Weight: 1.00
Name: AC_SE_SE94 Len: 232 Check: 3488 Weight: 1.00
Name: ACD_SE_SE8 Len: 232 Check: 3016 Weight: 1.00
Name: ACG_BE_VI1 Len: 232 Check: 5006 Weight: 1.00
Name: AD_SE_SE69 Len: 232 Check: 3362 Weight: 1.00
Name: AD_SE_SE71 Len: 232 Check: 2262 Weight: 1.00
Name: ADHK_NO_97 Len: 232 Check: 8765 Weight: 1.00
Name: ADK_CD_MAL Len: 232 Check: 6397 Weight: 1.00
Name: AG_BE_VI11 Len: 232 Check: 6471 Weight: 1.00
Name: AG_NG_92NG Len: 232 Check: 2880 Weight: 1.00
Name: AGHU_GA_VI Len: 232 Check: 9053 Weight: 1.00
Name: AGU_CD_Z32 Len: 232 Check: 523 Weight: 1.00
Name: AJ_BW_BW21 Len: 232 Check: 3842 Weight: 1.00
Name: B_AU_VH Len: 232 Check: 8468 Weight: 1.00
Name: B_CN_RL42 Len: 232 Check: 9366 Weight: 1.00
Name: B_DE_D31 Len: 232 Check: 3989 Weight: 1.00
Name: B_DE_HAN Len: 232 Check: 563 Weight: 1.00
Name: B_FR_HXB2 Len: 232 Check: 3184 Weight: 1.00
Name: B_GA_OYI Len: 232 Check: 5511 Weight: 1.00
Name: B_GB_CAM1 Len: 232 Check: 4779 Weight: 1.00
Name: B_GB_GB8 Len: 232 Check: 1128 Weight: 1.00
Name: B_GB_MANC Len: 232 Check: 2885 Weight: 1.00
Name: B_KR_WK Len: 232 Check: 9915 Weight: 1.00
Name: B_NL_3202A Len: 232 Check: 3135 Weight: 1.00
Name: B_TW_TWCYS Len: 232 Check: 2211 Weight: 1.00
Name: B_US_BC Len: 232 Check: 3145 Weight: 1.00
Name: B_US_DH123 Len: 232 Check: 7019 Weight: 1.00
Name: B_US_JRCSF Len: 232 Check: 4099 Weight: 1.00
Name: B_US_MNCG Len: 232 Check: 4137 Weight: 1.00
Name: B_US_P896 Len: 232 Check: 4405 Weight: 1.00
Name: B_US_RF Len: 232 Check: 450 Weight: 1.00
Name: B_US_SF2 Len: 232 Check: 5413 Weight: 1.00
Name: B_US_WEAU1 Len: 232 Check: 5335 Weight: 1.00
Name: B_US_WR27 Len: 232 Check: 3720 Weight: 1.00
Name: B_US_YU2 Len: 232 Check: 9943 Weight: 1.00
Name: BF1_BR_93B Len: 232 Check: 3598 Weight: 1.00
Name: C_BR_92BR0 Len: 232 Check: 3908 Weight: 1.00
Name: C_BW_96BW0 Len: 232 Check: 3880 Weight: 1.00
Name: C_BW_96BW1 Len: 232 Check: 4542 Weight: 1.00
Name: C_BW_96BW1 Len: 232 Check: 6118 Weight: 1.00
Name: C_BW_96BW1 Len: 232 Check: 5089 Weight: 1.00
Name: C_ET_ETH22 Len: 232 Check: 744 Weight: 1.00
Name: C_IN_93IN1 Len: 232 Check: 943 Weight: 1.00
Name: C_IN_93IN9 Len: 232 Check: 1241 Weight: 1.00
Name: C_IN_93IN9 Len: 232 Check: 9885 Weight: 1.00
Name: C_IN_94IN1 Len: 232 Check: 6448 Weight: 1.00
Name: C_IN_95IN2 Len: 232 Check: 5597 Weight: 1.00
Name: CRF01_AE_C Len: 232 Check: 1052 Weight: 1.00
Name: CRF01_AE_C Len: 232 Check: 744 Weight: 1.00
Name: CRF01_AE_C Len: 232 Check: 1265 Weight: 1.00
Name: CRF01_AE_T Len: 232 Check: 697 Weight: 1.00
Name: CRF01_AE_T Len: 232 Check: 8468 Weight: 1.00

```
Name: CRF01_AE_T      Len:   232  Check: 9246  Weight:   1.00
Name: CRF01_AE_T      Len:   232  Check: 8105  Weight:   1.00
Name: CRF01_AE_T      Len:   232  Check: 9948  Weight:   1.00
Name: CRF01_AE_T      Len:   232  Check: 9460  Weight:   1.00
Name: CRF02_AG_F      Len:   232  Check: 925  Weight:   1.00
Name: CRF02_AG_F      Len:   232  Check: 9559  Weight:   1.00
Name: CRF02_AG_G      Len:   232  Check: 399  Weight:   1.00
Name: CRF02_AG_N      Len:   232  Check: 2782  Weight:   1.00
Name: CRF02_AG_S      Len:   232  Check: 538  Weight:   1.00
Name: CRF02_AG_S      Len:   232  Check: 6700  Weight:   1.00
Name: CRF03_AB_R      Len:   232  Check: 6784  Weight:   1.00
Name: CRF03_AB_R      Len:   232  Check: 3106  Weight:   1.00
Name: CRF04_cpx_      Len:   232  Check: 1551  Weight:   1.00
Name: CRF04_cpx_      Len:   232  Check: 5866  Weight:   1.00
Name: CRF04_cpx_      Len:   232  Check: 7925  Weight:   1.00
Name: CRF05_DF_B      Len:   232  Check: 3625  Weight:   1.00
Name: CRF05_DF_B      Len:   232  Check: 5585  Weight:   1.00
Name: CRF06_cpx_      Len:   232  Check: 3770  Weight:   1.00
Name: CRF06_cpx_      Len:   232  Check: 4202  Weight:   1.00
Name: CRF06_cpx_      Len:   232  Check: 5376  Weight:   1.00
Name: CRF06_cpx_      Len:   232  Check: 1869  Weight:   1.00
Name: CRF11_cpx_      Len:   232  Check: 3479  Weight:   1.00
Name: CRF11_cpx_      Len:   232  Check: 3712  Weight:   1.00
Name: D_CD_84ZR0      Len:   232  Check: 1380  Weight:   1.00
Name: D_CD_ELI        Len:   232  Check: 4418  Weight:   1.00
Name: D_CD_NDK        Len:   232  Check: 4588  Weight:   1.00
Name: D_UG_94UG1      Len:   232  Check: 2178  Weight:   1.00
Name: F1_BE_VI85      Len:   232  Check: 4350  Weight:   1.00
Name: F1_BR_93BR      Len:   232  Check: 7703  Weight:   1.00
Name: F1_FI_FIN9      Len:   232  Check: 5036  Weight:   1.00
Name: F1_FR_MP41      Len:   232  Check: 84  Weight:   1.00
Name: F2_CM_MP25      Len:   232  Check: 2622  Weight:   1.00
Name: F2KU_BE_VI      Len:   232  Check: 2193  Weight:   1.00
Name: G_BE_DRCBL      Len:   232  Check: 2548  Weight:   1.00
Name: G_NG_92NG0      Len:   232  Check: 3608  Weight:   1.00
Name: G_SE_SE616      Len:   232  Check: 2716  Weight:   1.00
Name: H_BE_VI991      Len:   232  Check: 1561  Weight:   1.00
Name: H_BE_VI997      Len:   232  Check: 663  Weight:   1.00
Name: H_CF_90CF0      Len:   232  Check: 1804  Weight:   1.00
Name: J_SE_SE702      Len:   232  Check: 1615  Weight:   1.00
Name: J_SE_SE788      Len:   232  Check: 1704  Weight:   1.00
Name: K_CD_EQTB1      Len:   232  Check: 4783  Weight:   1.00
Name: K_CM_MP535      Len:   232  Check: 2033  Weight:   1.00
Name: N_CM_YBF30      Len:   232  Check: 6419  Weight:   1.00
Name: O_CM_ANT70      Len:   232  Check: 8742  Weight:   1.00
Name: O_CM_MVP51      Len:   232  Check: 5835  Weight:   1.00
Name: O_SN_MP129      Len:   232  Check: 8625  Weight:   1.00
Name: O_SN_MP130      Len:   232  Check: 8793  Weight:   1.00
Name: U_CD___83C      Len:   232  Check: 1586  Weight:   1.00

              1                                                  50
00BW0762_1    MGGKWSKSS. IVGWPAVRER IR....RTDP .......... ....AAEGVG
00BW0768_2    MGGKWSKSSI V.GWPEVRER IRR..TEP.. .......... ....AAEGVG
00BW0874_2    MGGKWSKSS. LTGWPAVRER IR....RTEP .......... ....AAEGVG
00BW1471_2    MGGKWSKSS. IVGWPAVKER IRR..TNPR. .......... .TERAAVGVG
00BW1616_2    MGNKWSKSS. IVGWPAVRDR MRR..AEP.. .......... ....AAEGVG
00BW1686_8    MGGKWSKRS. KADWPAVREK LR....TTEP .......... ....AAEGVG
00BW1759_3    MGNKWSKS.. ...WPAVRER IRR..TRPAR .......... GNEPAAEGVG
00BW1773_2    MGSKWSKSSI V.GWPKVRET IRR..TEP.. .......... ....AAEGVG
00BW1783_5    MGNKWSKS.. ...WPAIRER IRR..TNPAA .......... ERTRAAEGVG
00BW1795_6    MGGKWSKSS. VVGWPAIRER MRR..... .......... .TEPAAEGVG
```

```
00BW1811_3  MGGKWSKSC. KIGWPAVRER MRR..... ..... .TEPAVEGVG
00BW1859_5  MGGKWSKSG. KVGWPEVRER MRR..TRPAA EGG..... ..DSAAEGVG
00BW1880_2  MGGKWSKSS. LVGWPAVRER IRT..TAP.. ..... .....
00BW1921_1  MGGKWSKSS. IVGWPAVRER MR....KTEP ..... ....AAEGVG
00BW2036_1  MGGKWSKSS. IVGWPAVRER IRR..... ..... .TEPAAEGVG
00BW2063_6  MGGKWSKSSI I.GWPAVRER MRK..AEP.. ..... ....AAEGVG
00BW2087_2  MGSKWSKSS. IVGWPAVRER IRR..T.... ..... RTEPAAEGVG
00BW2127_2  MGGKWSKSSI I.GWPAIRER IRR..TEP.. ..... ....AAEGVG
00BW2128_3  MGSKWSKCSI I.GWPAVRER IRR..AEP.. ..... ....AAVGVG
00BW2276_7  MGSKWSKC.. .SGWPDVRER MRR..ATPAA EAGRA...AP AAEGAAPGVG
00BW3819_3  MGSKWSKCSI V.GWPDVRER MRR..ARPAV RERRRQTEPA AEGVAAEGVG
00BW3842_8  MGGKWSKGR. IVGWPAVRER MRR..... ..... .TEPAAEGVG
00BW3871_3  MGSKWSKRS. IVEWPAVRER LR....KTEP ..... ....AAEGVG
00BW3876_9  MGGKWSKSS. IVGWPAVRER IRQ..TGAR. ..... ....AAEGVG
00BW3886_8  MGGKWSKSS. IVGWSAVRER MK....RTEP ..... ....AAEGVG
00BW3891_6  MGGKWSKSS. IVGWPTVRER MRR..TQP.. ..... ....AAEGVG
00BW3970_2  MGSKWSKRS. TAGWPAVRER MRR..TQPAA EG..... .TQSAAEGVG
00BW5031_1  MGGKWSKSS. LVGWPEVRDR IRR..TDP.. ..... ....AAEGVG
 96BW01B21  MGGKWSKSSI V.GWPAVRER IRR..TEP.. ..... ....AAEGVG
  96BW0407  MGGKWSKSSI V.GWPAVRER MRR..AEP.. ..... ....AAEGVG
  96BW0502  MGGKWSK... CSGWPAVRER MRR..TRPAV EGR..... .TESAAEGVG
 96BW06_J4  MGGKWSKSS. IVGWPAVRER IR....RTDP ..... ....PAEGVR
 96BW11_06  MGGKWSKSSI I.GWPAIRER IRR..TEPAA ER.....V ..GAAAEGVG
  96BW1210  MGNKWSKG.. ...WPAVRDR IRR..TEPAT ..... ..EPAAEGVG
 96BW15B03  MGGKWSKSS. IVGWPAVRER IRR..... ..... .TEPAAEGVG
 96BW16_26  MGGKWSK... ...WPAVRER MRR..TR... ..... .....VG
 96BW17A09  MGXKWSKRS. IVGWPNVRER IRR..TNPLT ER..... EAERAAVGVG
 96BWMO1_5  MGSKWSKSSI I.GWPAVRER IRK..TEPRK ..... .TEPAAEGVG
 96BWMO3_2  MGGKWSKSS. IVGWPAVRER MRR..TRPGA AE..... .....GVG
98BWMC12_2  MGSKWSKSS. IIGWPAVRER MR....RTEP ..... ....AAEGVG
98BWMC13_4  MGGKWSKSS. IIGWPAVRER MRR..... ..... .TEPAAEGVG
98BWMC14_a  MGGKWSKSS. LVGWPDVRER IR....KPRP KP..... ....AAEGVG
98BWMO14_1  MGSKLSKSK. IVGWPAIRER LR..... ..... RTEPAAEGVG
98BWMO18_d  MGGKWSKSS. IVGWPAVRER IRQ..TDPRE RI.....R QTEPAAEGVG
98BWMO36_a  MGGKWSKSSI V.GWPAVRER IRR..TEPRR ..... .AEPAAEGVG
98BWMO37_d  MGGKWSKSS. IVGWPEVRER LR....RTAP ..... ....AAEGVG
99BW3932_1  MGGKWSKRKI V.QWPTVRER LRR..TEP.. ..... .....AEGVG
99BW4642_4  MGGKWSKSS. IVGWPAVRER IRR..TQPAA EG..... .....VG
99BW4745_8  MGSKLSKSC. TAGWPTVRER IRQ..AEP.. ..... ....AAEGVG
99BW4754_7  MGGKWSKSS. IVGWANVRER MRR..... ..... .TEPAAVGVG
99BWMC16_8  MGNKWSKS.. ...WPAVRER IRR..TEPAV RV.....R RTEPAAEGVG
A2_CD_97CD  MGGKWSKRT. IVGWPEIRER MRR..TPPAA EG.....VR PTPPAAEGVG
A2_CY_94CY  MGGKWSKRS. IPGWPAIRER MRRTPPTAQR TE..... AVSPAAPGVG
A2D___97KR  MGGKWSKRS. LPGWPAIRER MRRTPPAAER TP..... PAA.AAEGVG
A2G_CD_97C  MGSKWSKSS. IVGWPAVRAR IR...QTPP. ..... ....AAEGVG
A_BY_97BL0  XXGKWSKSS. IXXWPQVXER IRRAPAP... ..... ....AARXVG
  A_KE_Q23  MGGKWSKSS. IVGWPEIRER MRRAPP.... ..... ....AAPGVG
A_SE_SE659  MGGKWSKSS. IVGWPEIRER MRRAPS.... ..... ...AAAPGVG
A_SE_SE725  MGSKWSKSS. IVGWREVRER LRQTLAAARG ..... .....VG
A_SE_SE753  MGGRWSKSR. IVGWPEVRER IRRAPP.... ..... ....AATGVG
A_SE_SE853  MGGKWSKRS. KEGWSEVREK IRQT..... ..... ..PPAAKGVG
A_SE_SE889  MGGKWSKSS. IVGWPKVRER MARTPP.... ..... ....AAKGVG
A_SE_UGSE8  MGNKWSK... ..GWPEVRER IRQARAPAHT ..... PAPTAATGVG
A_UG_92UG0  MGNKWSKSC. IVGWPEVRER IRQTPTAARE RTR..... QAPTAAKGVG
 A_UG_U455  MGGKWSKKS. RVEWPEVRKR MRETPA.... ..... ....AAKGVG
AC_IN_2130  MGGKWPKSS. VVGWPEVRER IRRTPA.... ..... ....AAPGVG
AC_RW_92RW  MGSKWSKCSP V.GWPAVRER LRQ..TEP.. ..... ....AAEGVG
AC_SE_SE94  MGGKWSKSS. IIGWPQIRER IRRTPP.... ..... ....AATGVG
ACD_SE_SE8  MGGKWLKSSI V.GWPAVRER IRR..TEP.. ..... ....AAEGVG
ACG_BE_VI1  MGGKWSKRS. KVEWPQVRER MRQ..TPIAA EA.....EG ...AAAEGVG
AD_SE_SE69  MGGKWSKSS. IVGWPAVRER IKR..T.... ..... ..DPAAEGVG
```

```
AD_SE_SE71  MGGKWSKSS. IVGWPEVRER MRRARAP... .......... ...SAAPGVG
ADHK_NO_97  MGGKWSKSS. IVGWPAIRER MRR..AEP.. .......... ....AAEGVG
ADK_CD_MAL  MGGKWSKSS. IVGWPKIRER IRRTPPTETG .......... ...VGAVSQD
AG_BE_VI11  MGGKWSKSS. PVGWSRVRER MRR..TPPAA EG........ ...AAAEGVG
AG_NG_92NG  IGGKWSKSS. IVGWPAVRER IR...QTP.. .......... ....PAEGVG
AGHU_GA_VI  MGGEWSRSS. IVGWSTIRER MRRAEP.... .......... ....AAAGVG
AGU_CD_Z32  MGNKWSKG.. ...WPAVRER IRQ..TPPAP P......... ....AAEGVG
AJ_BW_BW21  MGSNWSKS.S IIGWPQVRER MKR.....AP A........P ....AAEGVG
   B_AU_VH  MGGKGSKRI. RSEWPTVRER IIQ..AEPAA AG........ ........VG
  B_CN_RL42 MGGKWSKHS. MFGWPSVRER MKR..AEPAA DG........ ........VG
   B_DE_D31 MGGKWSKSS. VVGWPAIRER MK........ .......... RAEPAAEGVG
   B_DE_HAN MGGKWSK... CSGWPTVRER MKQAEP.... .......... ..EPAADGVG
  B_FR_HXB2 MGGKWSKSS. VIGWPTVRER MR........ .......... RAEPAADRVG
   B_GA_OYI MGGKWSKCS. MKGWPTIRER MKR..AELQP PE........ ...PAAEGVG
  B_GB_CAM1 MGGKWSKRS. LGGWSAVRER MQR..AEP.. .......... RAEPAAEGVG
   B_GB_GB8 MGGKLSKRS. MFGWSRVRDR MQQ..AEP.. .......... ....AAEGVG
  B_GB_MANC MGGKWSKSR. KIGWPTVRER MKQ..VDPAE EGR.....KK QAEPAAEGVG
    B_KR_WK MGGKWSKRS. VPGWNTIRKR MRR..AEPAA EG........ ........VG
 B_NL_3202A MGGKWSKSS. VVGWPAIRER MK........ .......... RAEPAADGVG
 B_TW_TWCYS MGGKWSKRS. IPGWSNIRER IRQ..AEPA. .......... ....AADGVG
    B_US_BC MGGKWSKRM. EGGWHAVRER MR........ .......... RAEPAADGVG
 B_US_DH123 MGGKLSKCG. GVGWSTVRER MRR..AEPAA DR........ ..EP.AVGVG
 B_US_JRCSF MGGKWSKHS. VPGWSTVRER MRR..AEPAT DR......VR QTEPAAVGVG
  B_US_MNCG MGGKWSKR.. VTGWPTVRER MRR..AEP.. .......... .AELAADGVG
  B_US_P896 MGGKWSKRR. AEGWQTIRER MRRAEPA... .......... ..EPAADGVG
    B_US_RF MGGKWSKSK. MGGWPAVRER MQK..AEPAA DG........ ........VG
   B_US_SF2 MGGKWSKRS. MGGWSAIRER MRR..AEP.. .......... RAEPAADGVG
 B_US_WEAU1 MGGIWSKRS. GSGWPAIRER MKR..AEPAA EG........ ........VG
  B_US_WR27 MGGKWSKRS. VGGWPAIRER MX........ .......... RAEPAAEGVG
   B_US_YU2 MGGKWSKRS. MAGWPTVRER MRR..AEPAA ER......MR RAEPAADGVG
 BF1_BR_93B MGSKWSKSS. IVGWPAIRER LRQ..TP... .......... ...PAAEGVG
 C_BR_92BR0 MGNKWSKCST V.GRPAIRER MRR...AP.. .......... ....AAEGVG
 C_BW_96BW0 MGGKWSKSSI V.GWPAVRER MRR..TEP.. .......... ....AAEGVG
 C_BW_96BW1 MGGKWSKRSK I.EWPTIRDR MRR..TEPAA EG.......V ..GAAAEGVG
 C_BW_96BW1 MGNKWSKG.. ...WPAVRDR IRR..TEPAT .......... ..EPAAEGVG
 C_BW_96BW1 MGGKWSKSS. IVGWPAVRER IRR..... .......... .TEPAAEGVG
 C_ET_ETH22 MGGTMSKCSP V.GWPAIRER IRR..AAP.. .......... ....AAEGVG
 C_IN_93IN1 MGGKWSKCSI V.GWPAIRER MRR..AEP.. .......... ....AAEGVG
 C_IN_93IN9 MGGKWSKCSI V.GWPDIRER MRR..TQP.. .......... ....AAEGVG
 C_IN_93IN9 MGGKWSKCSI V.GWPAVRER MRR..TEP.. .......... ....AAEGVG
 C_IN_94IN1 MGGKWSKCSI V.GWPEIRER MRR..TQP.. .......... ....AAEGVG
 C_IN_95IN2 MGGKWSKCSI V.GWPDIRER MRR..TEP.. .......... ....AADGVG
 CRF01_AE_C MGGKWSKN.R IVGWPQVRER IRR..TPAAA .......... ......EGVG
 CRF01_AE_C MGGKWSKSC. IVGWPQVRER IRQ..TPVAE E........R QTPAAAEGVG
 CRF01_AE_C MGNKWSKS.. ...WPQIRER IRQ..TPVAT .......... ......EGVG
 CRF01_AE_T MGGKWSKS.S IVGWPQVREK IKQ..TPPAA .......... ......EGVG
 CRF01_AE_T MGGKWSKS.S IVGWPQVREK IKQ..TPPAA .......... ......EGVG
 CRF01_AE_T MGGKWSKS.S IVGWPQVRER IKQ..TPPAA .......... ......EGVG
 CRF01_AE_T MGAKWSKRG. ...WPQVRER IRQ..TPPAA .......... ......EGVG
 CRF01_AE_T MGSKWSKS.S IVGWPQVREK IKQ..TPPAT .......... ......EGVG
 CRF01_AE_T MGNKWSKS.. ...WPRVRER IKQ..TPPAA .......... ......EGVG
 CRF02_AG_F MGGKWSKSS. IVGWPKVRER IR...QTPP. .......... ....AATGVG
 CRF02_AG_F MGGKWSKSS. LVGWPKVRER II...QTPP. .......... ....AATGVG
 CRF02_AG_G MGGKWSKSS. IVGWPQVRER IR...QTPT. .......... ....AAKGVG
 CRF02_AG_N MGGKWSKSS. IVGWPKVMKR MR...QTPT. .......... ....AATGVG
 CRF02_AG_S MGGKWSKSS. IVGWPQIRDR IR...QTPP. .......... ....AARGVG
 CRF02_AG_S MGGKWSKSS. LVGWPQVRER IRRTQPTPS. .......... ....AAIGVG
 CRF03_AB_R MGGKWSKSS. IVGWPQVRER IRRAPAP... .......... ....AARGVG
 CRF03_AB_R MGGKWSKSS. IVGWPQIRER IRRAPAP... .......... ....AARGVG
 CRF04_cpx_ MGGKWSKSS. IVGWPEIRER MRR..ARAEP ERMR....RA QAEPAAAGVG
 CRF04_cpx_ MGGKWSKSS. LVGWPAIRER MRR..ARAEP ........AA QAEPAAAGVG
```

```
CRF04_cpx_   MGNKWSKS.. ...WPAVRER MRR..ARAEP .........A RAEPAAVGVG
CRF05_DF_B   MGGKWSKSS. VVGWPAIREK MRR..TP... .......... ...PAAEGVG
CRF05_DF_B   MGGKWSKNR. IVGWPAIRER MRR..TPPAA .......... ..GAAAEGVG
CRF06_cpx_   MGNKWSK... ..GWSQVRER MRR..TPPTE R......... ....AAEGVG
CRF06_cpx_   MGSKWSKS.S IVGWPQVRER IRQ..TPPTE G......... ....AAKGVG
CRF06_cpx_   MGGKWSKS.S LVGWPQVRER IRQ..TPPTE G......... ....AAEGVG
CRF06_cpx_   MGGKWSKS.S IVGWPKVRER MRQ..TPPAA E........R QTPPAAEGVG
CRF11_cpx_   MGGKWSKS.S IVGWPEIRER LRR..... .........T PPAAAADGVG
CRF11_cpx_   MGGNWSKS.S IVGWPEIRER LRR..... .........T PPTAAAEGVG
D_CD_84ZR0   MGGKWSKSS. IVGWPAIRER IRK..TDPRE RR....... RPEPAADGVG
  D_CD_ELI   MGGKWSKSS. IVGWPAIRER IRR..T.... .......... ..NPAADGVG
  D_CD_NDK   MGGKWSKSS. LVGWPAIRER IRK..T.... .......... ..DPAADGVG
D_UG_94UG1   MGGKWSKSS. IVGWPAVRER MRR..T.... .......... ..EPAAEGVG
F1_BE_VI85   MGGKWSKSS. IVGWPAVGER MRQ..TP... .......... ...TAAEGVG
F1_BR_93BR   MGGKWSKSS. IVGWPAIRER MRR..TPPT. .......... ..PPAAEGVG
F1_FI_FIN9   MGGKWSKSS. IVGWPAIRER MRR..PP... .......... ..PAAAEGVG
F1_FR_MP41   MGGKWSKSS. IVGWPAVRER MRR..TP... .......... ...PAAEGVG
F2_CM_MP25   MGGKWSKSS. IVGWPAIRER IRR..TP... .......... ...VAAEGVG
F2KU_BE_VI   MGGKWSK... ..GWPSVRER IRR..TPPAA P......... ....AADGVG
G_BE_DRCBL   MGNKWSKRK. VAGWPEVRER LR...QHPA. .......... ....AAEGVG
G_NG_92NG0   MGGKWSKSS. IVGWPQIRER IR...QTPV. .......... ....AAEGVG
G_SE_SE616   MGGKWSKSS. IVGWPEVRER IR...NTPT. .......... ....AAEGVG
H_BE_VI991   MGGKWSKGC. ISGWPAVRER IRQ..TEP.. .......... ....AAEGVG
H_BE_VI997   MGGKWSKSS. IVGWPAVRER IRR..AQP.. .......... ....AADGVG
H_CF_90CF0   MGGKWSKSR. MGGWSTIRER MRR..AEP.. .......... ....VAEGVG
J_SE_SE702   MGNKWSKS.. ...WPQVRDR MRR..A..AP A........P ....AADGVG
J_SE_SE788   MGNKWSKS.. ...WPQVRER MRR.....AP A........P ....AADGVG
K_CD_EQTB1   MGGKWSKS.S IVGWSTVRER MR..... .......... KTPPAADGVG
K_CM_MP535   MGGKWSKS.S IVGWPAIRER MRR..ARPAA DR.......V GTQPAADGVG
N_CM_YBF30   MGKIWSKSS. LVGWPEIRER MRRQTQEP.. .......... .AVEPAVGAG
O_CM_ANT70   MGNALRKGK. FEGWAAVRER MRRTRTF... .........P ESEPCAPGVG
O_CM_MVP51   MGNAWSKSK. FAGWSEVRDR MRRSSS.... .........D PQQPCAPGVG
O_SN_MP129   MGNVLGKDI. FKGWSAVRER MRGTS..... .........P DPEPCAPGVG
O_SN_MP130   MGNVLGKDK. FKGWSAVRER MRKTS..... .........P EPEPCAPGVG
U_CD___83C   MGNKWSKQ.. ...WPAIRER MRR..ARPAA E........P ....AADGVG

             51                                                  100
00BW0762_1   AASQDLDRHG AITSSNTPTT NAACAWLEAQ EEDE...VGF PVRPQVPLRP
00BW0768_2   AASQDLDKYG ALTSSNTANT NAACAWLEAQ EKEE.E.VGF PVRPQVPLRP
00BW0874_2   EASQDLDKHG ALTTSNTAAN NADCAWLQAH EEEG..EVGF PVRPQVPLRP
00BW1471_2   AASQDLDKYG ALTANNTPAN NADCARLQAQ EEDN..EVGF PVRPQVPLRP
00BW1616_2   PASQDLGRHG ALTTSNTAHN NADCAWLEAQ EEEG.G.VGF PVRPQVPLRP
00BW1686_8   AASQDLAKHG ALTSSNTAHN NADCAWLEAQ EEEE...VGF PVRPQVPLRP
00BW1759_3   AASQDLDKYG ALTSSNTPTT NAACAWLEAQ EEAE.E.VGF PVRPQVPLRP
00BW1773_2   AASQDLDKYG ALTTSNTDTN NADCAWLKAQ EEEG.E.VGF PVRPQVPLRP
00BW1783_5   AASRDLDKHG ALTSSNTPAT NADCAWLEAQ EEEE.E.VGF PVRPQVPVRP
00BW1795_6   AASQDLDKYG ALTTSNTATN NADCAWLEAQ EEAD..DVGF PVRPQVPLRP
00BW1811_3   AASQDLDRHG ALTSSNTATN NADCAWLEAQ EEGE..EVGF PVRPQVPLRP
00BW1859_5   AASRDLARHG ALTSSNTATT NAACAWLEAQ EEAD..EVGS PVRPQVPLRP
00BW1880_2   .......... .......AAT NDACARLKEQ EEEG.E.VGF PVRPQVPVRP
00BW1921_1   AASRDLDKYG ALTSSNTVTT NAACAWLEAQ EKEG..GVGF PVRPQVPLRP
00BW2036_1   AASQDLDKHG ALTSNNTSST NDACAWLEAQ EEAD..DVGF PVRPQVPLRP
00BW2063_6   AASQDLDKHG ALTSSNTANN NADCAWLEAQ EEEE...VGF PVKPQVPLRP
00BW2087_2   AASQDLDKHG ALTPSNTAGN NSTCAWLQAQ EEEE..EVGF PVRPQVPLRP
00BW2127_2   AASQDLDKHG ALTSSNTATN NADCAWLQAQ EEES.EGVGF PVRPQVPLRP
00BW2128_3   AASQDLDKYG ALTSSNTDTN NAACAWLRAQ EEEE.E.VGF PVX.....TS
00BW2276_7   AASRDLDRHG ALTSSNTPTN NADCAWLQAQ EEEE.EEVGF PVRPQVPLRP
00BW3819_3   TASQDLAKHG ALTSSNTPAN NAACAWLQAQ EEEE.D.VGF PVRPQVPLRP
00BW3842_8   AASRDLDKHG ALTSSNTAEN NADCAWLEAQ EETE..EVGF PVRPQVPLRP
00BW3871_3   AASQDLDKHG ALTSSNTAQN NAACAWLEAQ EEAE.GEVGF PVRPQVPLRP
00BW3876_9   AASQDLDRHG ALTSSNTGAT NSDVAWLEAQ EEEG.E.VGF PVRPQVPLGP
```

```
00BW3886_8   AASQDLDKHG ALTSSNTPAN NADCAWLETQ EEAD..EVGF PVRPQVPLRP
00BW3891_6   AASQDLDKYG ALTSSNTADT NADCAWLQAQ EEEG.E.VGF PVRPQVPLRP
00BW3970_2   AASQDLDKHG ALTSSNTAAT NADCAWLEAQ EEDG..DVGF PVRPQVPLRP
00BW5031_1   AASQDLDKHG ALTNSNTAAT NKDCAWLEAQ EEEG.E.VGF PVRPQVPLRP
 96BW01B21   AASQDLDKYG ALTTSNTDPN NAACAWLRAQ EEEG.E.VGF PVRPQVPLRP
  96BW0407   TASRDLDKYG ALTTNNTPTN NAACAWLEPQ KEEE.E.VGF PVRPQVPLRP
  96BW0502   AASQDLDKYG ALTSSNTSTN NADCAWLEAQ EEEG..DVGF PVRPQVPLRP
 96BW06_J4   APSQDLDKHG ALTSSNTPAN NADVAWLEPQ EEEG..DVGF PVRPQVPLRP
 96BW11_06   AASQDLDKHG ALTSSNTAAN NADCAWLEAQ EEEE...AVGF PLRPQVPLKP
  96BW1210   SASKDLEKHG ALTTSNTAAN NPHCAWLEAQ EEEE...VGF PVRPQVPLRP
 96BW15B03   AASQDLDKYG AFTSSNTASN NADCAWLEAQ EEAD..EVGF PVRPQVPLRP
 96BW16_26   AEYQDLDKDR ALTSSNTDTN NAACAQ...Q EEEE.E.VGF PVRPQVPLRP
 96BW17A09   AASQDLDKHG ALTSSNTPAN NADCAWLQAQ EEDE..EVGF PVRPQVPLRP
 96BWMO1_5   AVSQDLDQHG AITSSNTATN .ADCACLEAQ EEEA.E.LGF PVRPQMPLKP
 96BWMO3_2   AASQDLDKYG ALTSSNTASN NADCAWLEAQ EEAN..EVGF PVRPQVPLRP
98BWMC12_2   AASQDLDKYG AMTTSNTAHN NGDCAWLQAQ EEEE..EVGF PVRPQVPLRP
98BWMC13_4   AASQDLDKHG ALTSSNTATT NADCAWLEAQ AEAE..KVGF PVRPQVPLRP
98BWMC14_a   AASQDLDKYG ALTSSNTATN NAACAWLEAQ EEAG..EVGF PVKPQVPLRP
98BWMO14_1   AASQDLDKHG ALTTSNTASN NAACAWLQAQ EEDN.GDVGF PVRPQVPLRP
98BWMO18_d   AASRDLERHG AITTSNTAAN NPACAWLEAQ EEDG.D.VGF PVRPQVPLRP
98BWMO36_a   AASQDLDKHG ALTTSNTASN NADCAWLEAQ EEEE.E.VGF PVKPQVPVRP
98BWMO37_d   AASQDLAKHG AITSSNTATT NADCAWLEAQ EEGE..EVGF PVRPQVPLRP
99BW3932_1   AASQDLANHG ALTTSNTATN NADCAWLQAQ EEEE...VGF PVRPQVPLRP
99BW4642_4   AASQDLDRHG AITSSNTAAT NADCAWLEAQ EEAD..EVGF PVRPQX...P
99BW4745_8   AASKDLDKHG ALTSSNTAGT NADCAWLQAQ EEEG.E.VGF PVRPQVPLRP
99BW4754_7   AASQDLDKYG ALTSSNTAAT NADCARLEAQ EETE..EVGF PVRPQVPLRP
99BWMC16_8   AASKDLEKHG ALTTSNTVHN NPDCAWLEAQ EEEE.E.VGF PVRPQVPLRP
A2_CD_97CD   AVSQDLARHG AVTSSNTAAN NPDCAWLEAQ EEE...EVGF PVRPQVPLRA
A2_CY_94CY   AVSQDLATHG AVTSSNTAAT NPDCAWVEAQ EEES..EVGF PVRPQVPLRP
A2D___97KR   AVSQDLATRG AVTINNTAAT NADSAWLEAQ EEEE..EVGF PVRPQVPLRP
A2G_CD_97C   AVSQDLARHG AITSSNTAAN NPDCAWLEAQ QEDS..EVGF PVRPQVPLRP
A_BY_97BL0   PVSQDXDKHG AVTSSNTAAN NADCAXLEAQ X...EXEVGF PVRPQVPLRP
  A_KE_Q23   AVSQDLDKHG AVTSKNINH. .PSYAWLEAQ E...DEDVGF PVRPQVPLRP
A_SE_SE659   AVSQDLAKHG AVTSSNINN. .PSCVWLEAQ E...EEEVGF PVRPQVPLRP
A_SE_SE725   AVSQDLEKYG AVTSSNVNH. .PSCAWLEAQ E...EEEVGF PVRPHVPLRP
A_SE_SE753   AVSQDLDKYG AITSSNINH. .PSCVWLEAQ E...DEGVGF PVRPQVPLRP
A_SE_SE853   AVSQDLDRHG AITSSNINH. .PSCTWLEAQ E...DEEVGF PVKPQVPLRP
A_SE_SE889   AVSQDLDKHG AVTSSNINH. .PSCAWVEAQ E...EEEVGF PVRPQVPLRP
A_SE_UGSE8   AVSQDLEKHG AITSSNINH. .PSCTWLEAQ AQE.DEEVGF PVRPQVPLRP
A_UG_92UG0   AVSQDLDKHG AVTSSNVNH. .PSCVWLEAQ E...EEEVGF PVRPQVPLRP
 A_UG_U455   AVSQDLDKYG AVTSSNTSST NASCAWLEAQ E...EGDVGF PVRPQVPLRP
AC_IN_2130   AVSQDLDKHG AVTSSNVNH. .PSCVWLDAQ E...EEEVGF PVRPQVPLRP
AC_RW_92RW   AASQDLDKYG ALTSSNTPSN NADCAWLAAQ EEEN.E.VGF PVRPQVPLRP
AC_SE_SE94   AVSQDLDKHG AITSSNINH. .PSNTWLAAQ E.E.EEEVGF PVRPQVPLRP
ACD_SE_SE8   AASQDLDKYG ALTSSNTVTN NPDCAWLEAQ KEEE.E.VGF PVRPQVPLRP
ACG_BE_VI1   AVSQDLARHG AVTSRNTSAT NADCAWLEAQ EDE...EVGF PVRPQVPLRP
AD_SE_SE69   AASRDLGKHG AITSSNTAQT NPDCAWLEAQ EED..GEVGF PVRPQVPLRP
AD_SE_SE71   AVSQDLDKHG AITSSNINH. .PSCVWLEAQ E...DEDVGF PVRPQVPLRP
ADHK_NO_97   AASRDLDRRG AIXINNLPSN NSDSAWLEAQ EE...EEVGF PVRPQVPLRP
ADK_CD_MAL   AVSQDLDKCG AAASSSPAAN NASCEPPEEE EE.....VGF PVRPQVPLRP
AG_BE_VI11   AVSRDLAKHG AITSSNTVAT NADCAWLEAQ KEG..EEVGF PVRPQVPLRP
AG_NG_92NG   AAPQDLARHG AITSSNTAQT NPDCAWLEAQ QENS..EVGF PVRQQVPLRP
AGHU_GA_VI   AVSQDLDKRG AITTNNTVKT NADCAWLEAQ EDE...EVGF PVRPQLPLRP
AGU_CD_Z32   AASQDLAKHG AISSSNTATN NPDCAWLEAQ EESE..EVGF PVRPQVPLRP
AJ_BW_BW21   AVSRDLDKHG AITSSNTPTT NADCAWLEAQ TEAE.DEVGF PVKPQVPLRP
   B_AU_VH   AASRDLEKHG AITSSN..IN NADCVWLQAQ EEE...EVGF PVRPQVPLRP
B_CN_RL42    AVSRDLEKHG AITSSNTAAT NAACAWLEAQ EEE...EVGF PVRPQVPLRP
 B_DE_D31    AVSRDLEKHG AITSSNTPAT NAACAWLEAQ EEE...EVGF PVKPQVPLRP
 B_DE_HAN    AASRDLKKHG AITSSNTATN NAACAWLEAQ EEE...EEGF PVRPQVPLRP
B_FR_HXB2    AASRDLEKHG AITSSNTAAT NAACAWLEAQ EEE...EVGF PVTPQVPLRP
 B_GA_OYI    AASRDLEKHG AITSSNTAAT NADCAWLEAQ EDE...EVGF PVRPQVPLRP
```

```
B_GB_CAM1   AVSRDLEKHG AITSSNTAAT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
 B_GB_GB8   AVSRDLEKHG AITSSNTAAT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
B_GB_MANC   AVSRDLEKHG AITSSNTAAT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
  B_KR_WK   AASRDLEQRG AITTSNTASN NAACAWQEAQ EEE...EVGF PVRPQVPLRP
B_NL_3202A  AVSRDLEKHG AITSSNTAAT NADCAWLEAQ EDE...EVGF PVRPQVPLRP
B_TW_TWCYS  AVSRDLEKHG AITSSNTAAT NADCAWLEAQ EEE...EVGF PVKPQVPLRP
  B_US_BC   AVSRDLEKHG AITSSNTAAN NADCAWLEAQ EEE...EVGF PVRPQVPLRP
B_US_DH123  AASRDLGKHG AITSSNTAAT NADCAWLEAQ QEE..EEVGF PVRPQIPLRP
B_US_JRCSF  AVSRDLEKHG AITSSNTAAT NADCAWLEAY EDE...EVGF PVRPQVPLRP
 B_US_MNCG  AASRDLEKHG ALTSSNTAAT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
 B_US_P896  AVSRDLARHG AITSSN..TN NADIAWLEAQ EEG...EVGF PVKPQVPLRP
  B_US_RF   AASRDLEKHG TITSSNTAAN NAACTWLEAQ EDED.EEVGF PVRPQVPLRP
 B_US_SF2   AVSRDLEKHG AITSSNTAAT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
B_US_WEAU1  AVSRDLAKHG AITSSN..EN NADCVWLKAQ EDE...EVGF PVRPQVPLRP
 B_US_WR27  AVSRDLEKHG AITNSNTAAT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
  B_US_YU2  AVSRDLERHG AITSSNTAAT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
BF1_BR_93B  AVSQDLERRG AITSSNTGAN NPDLAWLEAQ EEEE...VGF PVRPQVPLRP
C_BR_92BR0  PASQDSDKYG ALTSSSTPAN NADCAWLEAQ QEEE.E.VGF PVRPQVPLRP
C_BW_96BW0  TASQDLDKHG ALTTSNTAAN NAACAWLEAQ EEEE.E.VGF PVRPQVPLRP
C_BW_96BW1  AASQDLDKHG ALTSSNTAAN NADCAWLEAQ EEEE...VGF PVRPQVPLRP
C_BW_96BW1  SASKDLEKHG ALTTSNTAAN NPHCAWLEAQ EEEE.E.VGF PVRPQVPLRP
C_BW_96BW1  AASQDLDKYG AFTSSNTASN NADCAWLEAQ EEAD..EVGF PVRPQVPLRP
C_ET_ETH22  AASRDLDKYG ALTSSNTPAN NPDCAWLEAQ EEEE.E.VGF PVRPQVPLRP
C_IN_93IN1  AASQDLDKHG ALTSSNTDTT NADCAWLRAQ EEEG.E.VGF PVTPQVPLRP
C_IN_93IN9  AASQDLDKYG ALTSSNTDIT NPDCAWLKAQ EEEE.E.VGF PVRPQVPLRP
C_IN_93IN9  AASQDLDKYG ALTSSNTDTT NADCAWLRAQ EEEE.E.VGF PVRPQVPLRP
C_IN_94IN1  AASQDLDKYG ALTSSNTDTT NADCAWLRTQ EEEE.E.VGF PVRPQVPLRP
C_IN_95IN2  AASQDLDRYG ALTSSNTDTT NTE.......  .EEE.G.VGF PVRPQVPLRP
CRF01_AE_C  AASQDLDKHG AITSSNMNN.  .ADCAWLEAQ EE...EEVGF PVRPQVPLRP
CRF01_AE_C  AVSQDLDKHG AVTSSNINN.  .ADNVWLEAQ EE...EEVGF PVRPQVPLRP
CRF01_AE_C  AVSQDLDKHG AVTSSNMNS.  .AASVWLEAQ ED...EEVGF PVRPQVPTRP
CRF01_AE_T  TVSQDLDKHG AVTSSNMNN.  .DDCVWLGAQ EE...EEVGF PVRPQVPLRP
CRF01_AE_T  AVSQDLDKHG AVTSSNMNN.  .ADCEWLRAQ EE...EEVGF PVRPQVPLRP
CRF01_AE_T  AVSQDLDKHG AVTSSNMNN.  .ADCVWLRAQ EE...EGVGF PVRPQVPLRP
CRF01_AE_T  AVSRDLDKHG AVTSSNMNN.  .ADSVWLRAQ EED..EEVGF PVRPQVPLRP
CRF01_AE_T  AVSQDLDKHG AITSSNIDN.  .ADCVWLRAQ ED...EEVGF PVMPQVPLRP
CRF01_AE_T  AVSQDLDKHG AVTSTNMDN.  .ADSVWLRAQ EED..EEVGF PVEPQVPLRP
CRF02_AG_F  AASQDLDRHG AITSSNTADT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
CRF02_AG_F  AASQDLDRHG AITSSNTAET NADCAWLEAQ EDE...EVGF PVKPQVPLRP
CRF02_AG_G  AASQDLDKYG AITSSNTAFT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
CRF02_AG_N  AASQDLDRHG AITSSNTAQT NPDCAWLEAQ EDE...NVGF PVRPQVPLRP
CRF02_AG_S  AASQDLDRHG AITSSNTAAT NAACAWLEAQ EEE...EVGF PVRPQVPLRP
CRF02_AG_S  AVSQDLDRHG AITSRNTAHT NADCAWLEAQ EEE...EVGF PVRPQVPLRP
CRF03_AB_R  PVSQDLDKYG AVTSSNTAAN NADCAWLEAQ E...EEEVGF PVRPQVPLRP
CRF03_AB_R  PVSQDLDKYG AVTSSNTAAN NADCAWLEAQ K...EEEVGF PVRPQVPLRP
CRF04_cpx_  AVSQDLDKHG AITINNTAAT NPDKTWLEAQ EEEE..EVGF PVRPQVPLRP
CRF04_cpx_  AVSRDLDKHG AITSSNTAAT NPDKAWLEAQ EEDE..EVGF PVRPQVPLRP
CRF04_cpx_  AVSQDLDKYG AITTSNTAAT NPDKAWLEAQ EEEE..EVGF PVRPQVPLRP
CRF05_DF_B  AVFQDLDNRG AVTIRNTVAT NPDMAWVEAQ EEEEE..VGF PVKPQVPLRP
CRF05_DF_B  SVSQDLERRG AITSNNTAAT NADLAWVEAQ EEEE...VGF PVRPQVPLRP
CRF06_cpx_  AVSQDLDKHG AITSSNTATT NAACAWVEAQ TEE...EVGF PVRPQVPVRP
CRF06_cpx_  AVSQDLDKHG AITSNNTPTT NAACAWLEEQ TED...EVGF PVRPQVPLRP
CRF06_cpx_  AVSQDLDTHG AITSSNTVTT NAACAWLEAQ TED...EVGF PVRPQVPLRP
CRF06_cpx_  AASQDLAKHG AITSSNTAAT NADCAWLEAQ SEDN.EEVGF PVRPQVPLRP
CRF11_cpx_  AVSKDLEKFG AITSSNTAHT NDTCAWLEAQ ED...EEVGF PVRPQVPLRP
CRF11_cpx_  AASKDLERHG AITSSNTSQN NAACAWLEAQ ED...EGVGF PVRPQVPLRP
D_CD_84ZR0  AVSRDLERHG AITSSNTATT NAACAWVEAQ EED..EEVGF PVRPQVPLRP
  D_CD_ELI  AVSRDLEKHG AITSSNTAST NADCAWLEAQ EES..DEVGF PVRPQVPLRP
  D_CD_NDK  AVSRDLEKHG AITSSNTAST NDTCAWLEAQ EES..EEVGF PVRPQVPLRP
D_UG_94UG1  AASRDLEKHG AITSSNTAQT NDACAWLEAQ EE...EEVGF PVRPQVPLRP
F1_BE_VI85  AVSRDLDRRG AITSSNTRTT NPDLAWLEAQ EEEE...VGF PVRPQVPVRP
```

```
F1_BR_93BR  AVSQDLERRG AITSSNTRAN NPDLAWLEAQ EEDE...VGF PVRPQVPLRP
F1_FI_FIN9  AVSQDLERRG AITSSNTGAT NPDLAWLEAQ EDEE...VGF PVRPQVPLRP
F1_FR_MP41  AVSQDLDRRG AVTSSNTAAT NPDLAWLEAQ EDEE...VGF PVRPQVPLRP
F2_CM_MP25  AVSQDLDKRG AITNSNTGAT NADLAWLEAQ EDEE...VGF PVRPQVPLRP
F2KU_BE_VI  AVSQDLAKHG AITSSNTSST NPDCAWLEAQ EEE...EVGF PVRPQVPLRP
G_BE_DRCBL  AVSQDLDRHG AITSRNTAGT NPDCAWLEAQ EEDS..EVGF PVGPQVPLRP
G_NG_92NG0  AVSQDLARHG AITSSNTATN NPDCAWLEAQ EEDS..EVGF PVRPQVPLRP
G_SE_SE616  AVSQDLDRHG AITSSNTAAN NPDCAWLEAQ EEDS..DVGF PVRPQVPLRP
H_BE_VI991  AVSQDLDRRG AVTINNIASN NADSAWLEAQ EEE..EEVGF PVRPQVPLRP
H_BE_VI997  AVSRDLDRRG AVTINNTAAT NPDVAWLEAQ EEA..EEVGF PVRPQVPLRP
H_CF_90CF0  AVSRDLDRRG AVTINNTAST NRDAAWLEAQ EDG..EEVGF PVRPQVPLRP
J_SE_SE702  AVSQDLAKHG AITSSNTAAT NDDCAWLEAQ T..E.EEVGF PVRPQVPLRP
J_SE_SE788  AVSQDLAKHG AITSSNTAAT NADCAWLEAQ T..E.EEVGF PVRPQXPLRP
K_CD_EQTB1  AVSQDLDKHG AVTSSNTAFN NPDCAWLEAQ ED...EDVGF PVKPQIPLRP
K_CM_MP535  AVSQDLARHG AVTSSNTSHN NPDCAWLEAQ EE...EEVGF PVRPQVPLRP
N_CM_YBF30  AASQDLANRG AITIRNTRDN NESIAWLEAQ EEEE..EVGF PVRPQVPLRP
O_CM_ANT70  QISRELAARG GIPSSHTPQN NAALAFLESH QEEE...VGF PVAPQVPLRP
O_CM_MVP51  AVSRELATRG GISSSHTPQN NAALAFLDSH KDED...VGF PVRPQVPLRP
O_SN_MP129  QISRELAARG GIPSSYTPQN NAALAFLESH QDEE...VGF PVRPQVPLRP
O_SN_MP130  QVSRELAARG GISNSHTPQN NAALAFLESH QDED...VGF PVRPQVPLRP
U_CD___83C  AASRDLEKYG AITSSNTAET NEACAQLEAA QE.D.GEVGF PVRPQVPLRP

            101                                                 150
00BW0762_1  MTYKGALDLG FFLK.EKGGL EGLIYSTKRK EILDLWVYHT QGYFPDWQNY
00BW0768_2  MTYKGAVDLS WFLK.EKGGL EGLIYSKKRQ EILDLWVYNT QGYFPDWQNY
00BW0874_2  MTFKGAFDLG FFLK.EKGGL EGLIWSQKRQ DILDLWVYHT QGYFPDWQNY
00BW1471_2  MTYKGAFDLG FFLK.EKGGL DGMIYSKKRQ DILDLWVYNT QGYFPDWQNY
00BW1616_2  MTYKAAFDLS FFLK.EKGGL DGLIYSKKRK EILDLWVYHT QGFFPDWQCY
00BW1686_8  MTYKGAFDLG FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
00BW1759_3  MTYKAAFDLG FFLK.EKGGL EGLIHSKQRQ DILDLWVYHT QGYFPDWQNY
00BW1773_2  MTYKAAFDLS FLLK.EXGGL EGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
00BW1783_5  MTYKAAFDLS FFLK.EKGGL EGLIHSKQRQ DILDLWVYNT QGYFPDWQNY
00BW1795_6  MTYKAAFDLS FFLK.EKGGL DGLIYSRKRQ EILDLWVYHT QGFFPDWQNY
00BW1811_3  MTYKAAFDLS FFLK.EKGGL DGLIHSKKRQ DILDLWVYHT QGYFPDWQNY
00BW1859_5  MTYKAAFDLG FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
00BW1880_2  MTYKAAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGFFSDWQNY
00BW1921_1  MTYKAAFDLG FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
00BW2036_1  MTYKAAIDLS FFLK.EKGGL DGLIYSAKRQ EILDLWVYHT QGFFPDWQNY
00BW2063_6  MTMKGAVDLS FFLK.EKGGL EGLIYSKRRQ EILDLWVYHT QGFFPDWQNY
00BW2087_2  MTYKAAFDLS FFLK.EKGGL DGLIYSKQRQ DILDLWVYNT QGFFPDWQNY
00BW2127_2  MTYKAAFDLS FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
00BW2128_3  GTFKAAFDLS FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
00BW2276_7  MTYKAAFDLS FFLK.EKGGL EGLIYSRKRQ EILDLWVYHT QGYFPDWQNY
00BW3819_3  MTYKGAFDLS FFLK.EKGGL EGLIYSKKRQ DILDLWVYNT QGYFPDWQSY
00BW3842_8  MTYKGAVDLS FFLK.EKGGL DGLIYSQKRQ DILDLWVHHT QGYFPDWQNY
00BW3871_3  MTYKGAFDLG FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
00BW3876_9  MTYKGAVDLS FFLK.EKGGL ERLIYSKKRQ EILDLWVYHT QGYFPDWQNY
00BW3886_8  MTYKGAFDLG FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
00BW3891_6  MTYKGAVDLS FFLKXEKGGL EGLIYSKKRQ EILDLWVYHI QGYFPDWQNY
00BW3970_2  MTYKGAFNLG FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
00BW5031_1  MTYKGAFDLG WFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
 96BW01B21  MTYKGAFDLG FFLK.EKGGL EGLIYSKKRQ EILDLWVYNT QGVFPDWQNY
  96BW0407  MTYKAAVDLS FFLK.EKGGL GGIIYSNKRQ DILDLWVYNT QGYFPDWQNY
  96BW0502  MTYKGAFDLG FFLK.EKGGL EGLVYSKKRQ EILDLWVYHT QGFFPDWQNY
 96BW06_J4  MTYKGAVDLG FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
 96BW11_06  MTYKAAFDLS FFLK.EKGGL EGLIYSRKRQ EILDLWVYHT QGFFPDWHNY
  96BW1210  MTYKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGYFPDWQCY
 96BW15B03  MTYKGAVDLS FFLK.EKGGL DGLIYSPKRQ EILDLWVYHT QGFFPDWHNY
 96BW16_26  MTYKEAFDLS FFLK.EKGGL DGLIYSKKRQ DILDLWVYNT QGFLPDWQNY
 96BW17A09  MTYKAAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
 96BWMO1_5  VTYKAAFDLS FFLK.KKGGL EGLIYSKKRQ NILDLWVYHT QGFFPDWHNY
```

```
 96BWMO3_2  MTQKGAFDLG FFLK.EKGGL DGLIYSRKRQ EILDLWVYHT QGYFPDWQNY
98BWMC12_2  MTFKGALDLG FFLK.EKGGL DGLIHSKRRQ DILDLWVYHT QGYFPDWQNY
98BWMC13_4  MTYKGAVDLS FFLK.EKGGL DGLIYSKQRQ DILDLWVYNT QGYFPDWQNY
98BWMC14_a  MTYKAAFDLG FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
98BWMO14_1  MTYKAAFDLG FFLK.EKGGL DGLIYSKKRQ DILDLWVYNT QGFFPDWQNY
98BWMO18_d  MTYKAAFDLS FFLK.EKGGL DGLIYSKKRQ DILDLWVYNT QGFFPDWQNY
98BWMO36_a  ITYKAAFDLS FFLK.EKGGL EGLIYSKKRQ EILDLWVYNT QGYFPDWQNY
98BWMO37_d  MTYKGAFDLG FFLK.EKGGL DGLIHSKKRQ DILDLWVYHT QGYFPDWQNY
99BW3932_1  MTYKAAFDLS FFLK.EKGGL EGLIYSKQRQ DILDLWVYNT QGFFPDWHNY
99BW4642_4  MTYKAAFDLS FFLK.EKGGL DGLIYSRKRQ DILDLWVYHT QGFFPDWQNY
99BW4745_8  MTYKSAFDLS FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
99BW4754_7  MTYKAAFDLS FFLK.EKGGL DGLIHSKKRQ EILDLWVHHT QGFFPDWQNY
99BWMC16_8  MTYKAAVDLS WFLK.EXGGL DGLIHSQKRQ EILDLWVYYT QGYFPDWQNY
A2_CD_97CD  MTYKGAVDLS HFLK.EKGGL DGLIYSQRRQ DILDLWVYNT QGYFPDWQNY
A2_CY_94CY  MTFKGAFDLS FFLK.EKGGL DGLIYSQKRQ DILDMWVYHT QGYFPDWQNY
A2D___97KR  MTYKGAFDLS HFLR.EKGGL DGLIHSQKRQ DILDLWVYHT QGFFPDWQGY
A2G_CD_97C  MTYKSAFDLS FFLK.EKGGL DGLIYSKQRQ DILDLWVYNT QGFFPDWQNY
A_BY_97BL0  MTXXXXXDXS HFXK.EKGGL DGXIYSKKRQ XILDLXVYHT QGYFPDWQNY
  A_KE_Q23  MTYKGAVDLS HFLK.KKGGL DGLVYSRKRQ EILDLWVYHT QGYFPDWQNY
A_SE_SE659  MTYKAAVDLS HFLK.EKGGL DGLIYSRKRQ EILDLWVHHI QGFFPDWQNY
A_SE_SE725  MTYKGALDLS HFLK.EKGGL DGLIYSRRRQ EILDLWVYNT QGYFPDWQNY
A_SE_SE753  MTYKGALDLS HFLK.EKGGL DGLVYSKKRQ EILDLWVHHT QGYFPDWQNY
A_SE_SE853  MTYKGAFDLS HFLK.EKGGL DGLIYSRKRQ EILDLWVYNT QXYFPDWQNY
A_SE_SE889  MTYKGAVDLS HFLK.EKGGL DGLIYSRRRQ EILDLWVYNT QGYFPDWQNY
A_SE_UGSE8  MTYKGALDLS HFLK.EKGGL EGLIYSRKRQ EILDLWVYHT QGYFPDWHNY
A_UG_92UG0  MTYKAAFDLG FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
 A_UG_U455  MTYKAAFDLS FFLK.EKGGL DGLIHSQKRQ EILDLWVYHT QGFFPDWQNY
AC_IN_2130  MTYKGALDLS FFLK.EKGGL DGLIYSRKRQ EILDLWVYHT QGFFPDWHNY
AC_RW_92RW  MTYKAAVDLS FFLK.EKGGL EGLIYSKKRQ DILDLWVYNT QGYFPDWQNY
AC_SE_SE94  MTYKGALDLS HFLK.EKGGL DGLIYSQQRQ DILDLWVYNT QGYFPDWQNY
ACD_SE_SE8  MTYKAAVDLS FFLK.EKGGP EGLIYSKKRQ DILDLWVYNT QGFFPDWQNY
ACG_BE_VI1  MTFKGALDLS HFLK.EKGGL DGLIYSRKRQ EILDLWVYHT QGYFPDWQNY
AD_SE_SE69  MTYKGAVDLS HFLK.EKGGL EGLVWSPKRQ EILDLWVYHT QGYFPDWQNY
AD_SE_SE71  MTYKAAVDLS HFLK.EKGGL DGLIYSRKRA DILDLWVYHT QGFFPDWQCY
ADHK_NO_97  MTYKAAVDLS HFLK.EKGGL DGLIYSKKRQ EILDLWVHNT QGFFPDWQNY
ADK_CD_MAL  MTYKGAFDLS HFLK.EKGGL DGLVWSPKRQ EILDLWVYHT QGYFPDWQNY
AG_BE_VI11  MTYKAALDLS HFLK.EKGGL EGLIYSQQRQ DILDLWVYNT QGYFPDWQNY
AG_NG_92NG  MTYKAAFDLS FFLK.EKGGL DGLIYSKKRQ DILDLWVYNT QGYFPDWQNY
AGHU_GA_VI  MTFKGAFDLG FFLK.EKGGL DGLIHSQKRQ DILDLWVYHT QGYFPDWQNY
AGU_CD_Z32  MTFKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWHNY
AJ_BW_BW21  MTYKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGFFPDWQNY
   B_AU_VH  MTFKAAHDLS FFLK.EQGGL EGLIYSQRRQ DILDLWIYHT QGYFPDWQNY
 B_CN_RL42  MTYKGALDLS HFLR.EKGGL EGLIYSQRRQ DILDLWVYHT QGYFPDWQNY
  B_DE_D31  MTYKAAVDLS HFLK.EKGGL EGLVHSQKRQ DILDLWVYHT QGYFPDWQNY
  B_DE_HAN  MTYKGALDLS HFLK.EKGGL EGLIYSPKRQ EILDLWVYHT QGYFPDWQNY
 B_FR_HXB2  MTYKAAVDLS HFLK.EKGGL EGLIHSQRRQ DILDLWIYHT QGYFPD.QNY
  B_GA_OYI  MTYKGALDLS HFLK.EKGGL EGLIYSQKRQ DILDLWVYHT QGYFPDWQNY
 B_GB_CAM1  MTYKAALDIS HFLK.EKGGL EGIIYSQRRQ DILDLWIYHT QGYFPDWQNY
  B_GB_GB8  MTYKAAVDLS HFLK.EQGGL DGLIYSPKRQ EILDLWVYHT QGYFPDWQNY
 B_GB_MANC  MTFKGALDLS HFLR.EKGGL EGLVYSQKRQ DILDLWVYHT QGYFPDWQNY
   B_KR_WK  MTYKSALDLS HFLK.EKGGL EGLVYSQKRQ DILDLWVYHT QGFFPDWQNY
B_NL_3202A  MTYKGALDLS HFLK.EKGGL EGLIYSQRRQ DILDLWIYHT QGYFPDWQNY
B_TW_TWCYS  MDYKGAIDLS HFLR.KEGGL EGLVYSQKKE DILDLWIYHT QGFFPDWQNY
   B_US_BC  ITYKAAVDIS HFLK.EKGGL EGLIFSQRRQ DILDLWTYHT QGYFPDWQNY
B_US_DH123  MTYKAALDLS HFLK.EKGGL EGLIYSQKRQ DILDLWVYNT QGYFPDWQNY
B_US_JRCSF  MTYKAAIDLS HFLK.EKGGL EGLIYSQKRQ DILDLWIYHT QGYFPDWQNY
 B_US_MNCG  MTYKAALDLS HFLK.EKGGL DGLIYSQKRQ DILDLWVYHT QGYFPDWQNY
 B_US_P896  MTYKAAVDLS HFLK.EKGGL EGLVHSQKRQ DILDLWVYHT QGFFPDWQNY
   B_US_RF  MTFKAAVDLS HFLK.EKGGL DGLVFSQKRQ DILDLWVYHT QGYFPDWQNY
  B_US_SF2  MTYKAALDIS HFLK.EKGGL EGLIWSQRRQ EILDLWIYHT QGYFPDWQNY
B_US_WEAU1  MTYKAAHDLS HFK..EKGGL EGLIYSQKRQ DILDLWVYHT QGFFPDWQNY
```

```
B_US_WR27   MTYKAAVDLS HFLK.EKGGL EGLIYSQKRQ DILDLWVYHT QGYFPDWQNY
  B_US_YU2  MTHKAAMDLS HFLK.EKGGL EGLIHSQQRQ DILDLWVYHT QGYFPDWQNY
BF1_BR_93B  MTYKGALDLS HFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
C_BR_92BR0  MTYKAVVDLS FFLE.EKGGL EGLIYSKKRQ DILDLWVYNT QGYFPDWQNY
C_BW_96BW0  MTYKAAVDLS FFLK.EKGGL EGIIYSKKRQ DILDLWVYNT QGYFPDWQNY
C_BW_96BW1  MTYKAAFGLS FFLK.EKGGL EGLIYSRKRQ EILDLWVYHT QGYFPDWQNY
C_BW_96BW1  MTYKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGFFPDWHNY
C_BW_96BW1  MTYKGAVDLS FFLK.EKGGL DGLIYSPKRQ EILDLWVYHT QGYFPDWQCY
C_ET_ETH22  MTYKAAFDLS LFLK.EKGGL EGLIYSKKRQ EILDLWVYNT QGFFPDWHNY
C_IN_93IN1  MTYKSAFDLS FFLK.EKGGL EGLIYSKKRQ EILDLWVYNT QGFFPDWQNY
C_IN_93IN9  MTFKEAVDLS FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
C_IN_93IN9  MTFKGAFDLS FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
C_IN_94IN1  MTFKGAFDLS FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
C_IN_95IN2  MTFKGALDLS FFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
CRF01_AE_C  MTYKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
CRF01_AE_C  MTYKGAFDLS FFLK.EKGGL DGLIHSKRRQ EILDLWVHNT QGYFPDWQNY
CRF01_AE_C  MTYKGAFDLG FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGFFPDWQNY
CRF01_AE_T  MTYKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGFFPDWHNY
CRF01_AE_T  MTFKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGFFPDWQNY
CRF01_AE_T  MTYKGAFDLS FFLK.EKGGL EGLVYSKKRQ EILDLWVYHT QGFFPDWHNY
CRF01_AE_T  MTFKEAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGFFPDWQNY
CRF01_AE_T  MTYKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGFFPDWQNY
CRF01_AE_T  MTYKGAFDLS FFLE.EKGGL DGLVYSKKRQ EILDLWVYNT QGFFPDWQNY
CRF02_AG_F  MTYKAAFDLG FFLK.EKGGL DGLVYSKKRQ EILDLWVYNT QGFFPDWQNY
CRF02_AG_F  MTYKGAFDLG FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
CRF02_AG_G  MTYKGALDLS HFLK.EKGGL EGLIYSRKRQ EILDLWVYHT QGFFPDWQNY
CRF02_AG_N  MTYKGAVDLS HFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWHNY
CRF02_AG_S  MTYKAAVDLS HFLK.EKGGL EGLIYSKKRQ EILDLWVYNT QGFFPDWQNY
CRF02_AG_S  MTYKAAVDLS HFLK.EQGGL EGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
CRF03_AB_R  MTYKGAFDLS HFLK..KGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
CRF03_AB_R  MTYKGAFDLS HFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGYFPDWQ..
CRF04_cpx_  MTFKGALDLS HFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
CRF04_cpx_  MTFKAALDLS HFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWDNY
CRF04_cpx_  MTYKAAVDLS HFLK.EKGGL DGLIYSKQRQ DILDLWVYNT QGYFPDWQNY
CRF05_DF_B  MTYKAAVDLS HFLK.EKGGL EGLIYSKKRQ EILDLWVYNT QGYFPDWQNY
CRF05_DF_B  MTFKGALDLS HFLK.EKGGL DGLIWSRKRQ KILDLWVYNT QGYSPDWQNY
CRF06_cpx_  MTYKAAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGYFPDWQNY
CRF06_cpx_  MTFKGAFDLG FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGFFPDWQNY
CRF06_cpx_  MTFKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGYFPDWQNY
CRF06_cpx_  MTYKAAFDLG FFLK.EKGGL DGLIYSKKRQ DILDLWVYHT QGFFPDWQNY
CRF11_cpx_  MTYKAAFDLG FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
CRF11_cpx_  MTYKGAFDLS FFLK.EKGGL DGLIYSQKRK EILDLWVYHT QGFFPDWQEY
D_CD_84ZR0  MTYKAAVDLS HFLR.KKGGL EGLVYSQKRQ DILDLWVYHT QGFFPDWQNY
  D_CD_ELI  MTYKEALDLS HFLK.EKGGL EGLIWSKKRQ EILDLWVYNT QGIFPDWQNY
  D_CD_NDK  MTYKEAVDLS HFLK.EKGGL EGLIWSKKRQ EILDLWVYNT QGIFPDWQNY
D_UG_94UG1  MTYKEAVDLS HFLK.EKGGL EGLVWSPKRQ EILDLWVYHT QGFFPDWQNY
F1_BE_VI85  MTYKAAVDLS HFLK.EKGGL EGLIYSKKRG DTLDLWVYHT QGYFPDWQNY
F1_BR_93BR  MTYKGAVDLS HFLK.EKGGL EGLIYSKRRQ EILDLWVYHT QGYFPDWQNY
F1_FI_FIN9  MTYKGAFDLS QFXX.EKGGL EGLXYSKKRQ EILDLWLYHT QGYFPDWQNY
F1_FR_MP41  MTFKAAVDLS HFLK.EKGGL EGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
F2_CM_MP25  MTYKAALDLS HFLK.EKGGL EGLIYSRKRQ EILDLWVYHT QGYFPDWQNY
F2KU_BE_VI  MTYKGAFDLS HFIK.EKGGL DGLIYSKKRQ EILDLWVYNT QGYFPDWHNY
G_BE_DRCBL  MTYKAAFDLS FFLK.EKGGL DGLVYSKKRQ EILDLWVYHT QGFFPDWQNY
G_NG_92NG0  MTYKAAFDLS FFLK.EKGGL DGLIYSKRRQ DILDLWVYNT QGFFPDWQNY
G_SE_SE616  MTFKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGYFPDWQNY
H_BE_VI991  MTYKGAFDLS HFLK.EKGGL EGLIYSKKRQ EILDLWVYNT QGYFPDWHNY
H_BE_VI997  MTYKAALDLS HFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGYFPDWQNY
H_CF_90CF0  MTYKGAFDLS HFLK.EKGGL DGLIYSKQRQ DILDLWVYNT QGYFPDWQNY
J_SE_SE702  MTYKGAFDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVHNT QGYFPDWQNY
J_SE_SE788  MTYKGAVDLS FFLK.EKGGL DGLIYSKKRQ EILDLWVHNT QGYFPDWQNY
K_CD_EQTB1  MTFKGAFDLG FFLK.EKGGL DGLIYSKRRQ EILDLWVYHT QGFFPDWQNY
```

```
K_CM_MP535   MTYKAAFDLG FFLK.EKGGL DGLIYSKKRQ EILDLWVYHT QGFFPDWQNY
N_CM_YBF30   ITYKQAFDLS FFLK.DKGGL EGLVWSRKRQ DILDLWMYHT QGILPDWHNY
O_CM_ANT70   MTYKGAFDLS FFLK.EKGGL EGLIYSHKRA EILDLWVYNT QGFFPDWQNY
O_CM_MVP51   MTFKAAFDLS FFLK.EKGGL DGLIYSHKRA EILDLWIYHT QGFFPDWQCY
O_SN_MP129   MTYKGAFDLS FFLK.EKGGL DGLIYSHKRA EILDLWVYHT QGFFPDWQGY
O_SN_MP130   MTYKGAFDLS FFLK.EKGGL DGLIYSPERA EILDLWVYHT QGFFPDWQNY
U_CD___83C   MTYKAAFDLG FFLK.EKGGL DGLIYSKKRQ EILDLWVYNT QGYFPDWQNY

             151                                                 200
00BW0762_1   TPGPGTRYSL TFGWPFKLVP VD.PKEVEEA NEGENNCLLH PMSQHRIDDD
00BW0768_2   TPGPGVRYPL TLGWCFKLVP VD.PREAEEE TKGENSCLLH PMSQHGMEDS
00BW0874_2   TPGPGVRYPL TFGWPFKLVP VD.PTEVEEA NEGENNCLLH PMSQHGMDDK
00BW1471_2   TPGPGVRYPL TFGWCYKLVP VD.PKEVDAN .EGENNCLLH PMSQHGMDDE
00BW1616_2   TPGPGTRLPL TFGWCFKLVP VD.PREVEEA NEGENNCLLH PMSQQGMEDE
00BW1686_8   TPGPGVRYPL TFGWPFKLVP VD.PREVQEA NKGENNCLLH PMSQHGADDE
00BW1759_3   TPGPGIRYPL TFGWCFKLVP VD.PKEVEEA NEGEDNCLLH PMSLHGMEDD
00BW1773_2   TPGPGVRFPL TFGWCFKLVP VD.PREVEEA NEGEDNCLLH PMSQHGMEDE
00BW1783_5   TPGPGVRYPL TFGWCFKLVP VD.PREVEEA NEGENNCLLH PMSQHGMEDA
00BW1795_6   TPGPGTRYPL TFGWCYKLVP VD.PREVEEA NEGENNCLLH PMSQHGIEDG
00BW1811_3   TPGPGVRYPL TFGWCYKLVP VD.PGEVEEA NKGENNCLLH PMSLHGAEDI
00BW1859_5   TPGPGVRYPL TFGWCYKLVP VD.PGEVEEA NEGEDNCLLH PISQHGMEDT
00BW1880_2   TPGPGIRYPL TFGWCYKLVP VD.PGEVEEA NKGENNCLLH PMSQYGMDDE
00BW1921_1   TPGPGVRYPL TFGWPFKLVP VD.PKEVEEA NKGENNCLLH PLSQQGMEDE
00BW2036_1   TPGPGVRYPL TFGWCYKLVP VD.PREVEEA NQGENNCLLH PMSQFGMDDS
00BW2063_6   TPGPGVRYPL TFGWCFKLVP VD.PREVEEA NEGENNCLLH PISLHGMEDN
00BW2087_2   TPGPGVRYPL TSGWCYKLVP VD.PREVEEA NEGESNSLLH PISLHGMEDD
00BW2127_2   TPGPGVKYPL TFGWCFKLEP VD.PKEVEEA NEGDNNCLLH PMSQHGMDDE
00BW2128_3   TPGPGVRYPL TFGWRYKLVP VD.PQEVQQS NEGDNNCLLH PMSQHGIEDE
00BW2276_7   TPGPGIRYPL TFGWCFKLVP VD.PREVEEA NEGENNCLLH PRSLHGIEDE
00BW3819_3   TPGPGIRYPL TFGWCFKLVP VD.PREVEEN TEGENTCLLH PMSQHGMEDE
00BW3842_8   TPGPGVRYPL TFGWCYKLVP VD.PREVEEA NQGENNCLLH PMSQHGVDDP
00BW3871_3   TPGPGVRYPL TFGWPFKLVP VD.PREVEED NNGENNCLLH PMSQHGMDDE
00BW3876_9   TPGPGTRFPL TFGWCFMLVP VD.PREVEEA NEGENSCLLH PISQHGMEDE
00BW3886_8   TPGPGVRYPL TFGWPFKLVP VD.PREVEEA NQGENNCLLH PMSQYGMEDE
00BW3891_6   TPGPGVRYPL TFGWCFKLVP VD.PREVEEA NEGENNCLLH PISQHGMEDE
00BW3970_2   TPGPGVRYPL TFGWCFKLVP VD.PREIEEA NKGEDNCLLH PMSQHGMEDE
00BW5031_1   TPGPGVRYPL TFGWCYKLVP VD.PKEVEEA TEGENNCLLH PICQHGMEDE
 96BW01B21   TPGPGVRYPL TFGWCFKLVP VD.PREVEEA SEGEDNCLLH PMSQHGMEDE
  96BW0407   TPGPGVRYPL TFGWCYKLVP VD.PREVEEA NEGENNRLLH PMSLHGMEDD
  96BW0502   TPGPGVKYPL TFGWCFKLVP VD.PGEVEEA NKGEDNCLLH PMSQHGMEDE
 96BW06_J4   TPGPGITYPL TFGWPFKLVP VD.PREVEEA NNGENNCLLH PMSQHGMDDE
 96BW11_06   TPGPGVRYPL TFGWCFKLVP VD.PGEVEEA NEGENNCLLH PMAQHGMEDE
  96BW1210   TPGPGVRYPL TFGWCFKLVP VD.PGEVEEA NGGEDNCLLH PMSQHGIEDA
 96BW15B03   TPGPGVRYPL TFGWCYKLVP VD.PREVEEN NQGENHCLLH PMSQHGMDNP
 96BW16_26   TPGPGVRFPL TFGWCFKLVP VD.PREVEEE TEGDNNCLLH PMNQHGMDDP
 96BW17A09   TPGPGVRYPL TFGWCYKLVP VD.PREVEED NEGENNCLLH PMSQHGMEDA
 96BWMO1_5   TPGPGIRYPL TFGWCFKLVP VD.PREVEEA NKGENTSLLH PVSLHGMEDE
 96BWMO3_2   TSGPGVRYPL TFGWCYKLLP VD.PKEVEEA SEGENNCLLH PMNQHGMEDE
98BWMC12_2   TPGPGVRYPL TFGWPFKLVP VD.PKEVEEA NKGENNCLLH PMSLHGIEDA
98BWMC13_4   TPGPGVRYPL TFGWCYKLVP VD.PREVEEA NEGENNCLLH PMSLHGIEDA
98BWMC14_a   TPGPGTRYPL TFGWPFKLVP VD.PREVEEA NNGENNCLLH PMSQHGMDDE
98BWMO14_1   TPGPGVRYPL TFGWCYKLVP VD.PREVEEA NEGEDNCLLH PMSQHGIEDP
98BWMO18_d   TPGPGVRYPL TFGWCYKLVP VD.PREVEEA NEGEDNCLLH PMSQHGMEDA
98BWMO36_a   TPGPGTRYPI TFGWCFKLVP VD.PREVEEA NEGENNSLLH PVSLHGMEDE
98BWMO37_d   TPGPGVRYPL TFGWCFKLVP VD.PREVEEA TEGEDNCLLH PLSQHGIEDA
99BW3932_1   TPGPGVRYPL TFGWCFKLVP VD.PREVEEN NQRENSCLLH PMSQHGMEDP
99BW4642_4   TPGPGVRYPL TFGWCFKLVP VD.PREVEEA NEGENNCLLH PMSQHGMEDA
99BW4745_8   TPGPGVRYPL TFGWCLKLVP VD.PREVEED NEGENNCLLH PMSQHGMEDA
99BW4754_7   TPGPGIRYPL TFGWCFKLVP VD.PREVQEA NEGDDNCLLH PMSQFGMDDD
99BWMC16_8   TPGPGIRYPL TFGWCFRLVP VD.PSEVEEA NQGENNCLLH PISQHGMEDE
A2_CD_97CD   TPGPGARFPL TFGWCFKLVP VD.PSEVEEA TEGENNSLLH PICQHGAEDP
```

```
A2_CY_94CY   TPGPGIRYPL TFGWCFKLVP VE.PSEVEEA TQGENNSLLH PICQHGVDDP
A2D___97KR   TPGPGIRYPL TFGWCFKLVP VS.PAEVEEA TEGETNSLLH PICQHGIEDP
A2G_CD_97C   TPGPGVRYPL TFGWCFKLVP MD.PAEVEEA NKEENSSLLH PICQHGMEDD
A_BY_97BL0   TPGPXIRFPL TFXXCYKLVP VD.PAEVEEA TXGENNSLLH PICQHGMDDE
  A_KE_Q23   TPGPGTRFPL TFGWCFKLVP VD.PDEVEKA TEGENNSLLH PICQHGMDDE
A_SE_SE659   TPGPGIRFPL TFGWCFKLVP VD.PDEVEKD TEGENNSLLH PICQHGMDDE
A_SE_SE725   TPGPGVRYPL TFGWCFKLVP VD.PDEVEQA NEGENNSLLH PMCQHGMDDE
A_SE_SE753   TPGPGIRYPL TFGWCFKLVP VD.PDEVKKD TEGENNSLLH PMCQHGMDDE
A_SE_SE853   TPGPGVRYPL TFGWCFKLVP VE.PEEVEKA NEGENNSLLH PICQHGMDDE
A_SE_SE889   TPGPGTRFPL TFGWCFKLVP VD.PDEVEKA TEGENNSLLH PICQHGMDDE
A_SE_UGSE8   TPGPGIRYPL TFGWCFKLVP VD.PDEVEKA TEGENNSLLH PMCQHGMDDE
A_UG_92UG0   TPGPGIRYPL TFGWCFKLVP VD.EDEVEEA TGGENNSLLH PICQHGMDDE
 A_UG_U455   TPGPGIRYPL TFGWCYKLVP VD.PAEVEEA TGGENNSLLH PICQHGVDDE
AC_IN_2130   TPGPGVRYPL TFGWCFKLVP VD.PREVEEA TGGEDNCLLH PVCQHGMEDE
AC_RW_92RW   TPGPGVRYPL TFGWCYKLVP VD.PREVEEA NEGEDNCLLH PLSQHGMEDE
AC_SE_SE94   TPGPGIRYPL TFGWCFKLVP VN.PDEVEEA TKGENNSLLH PMCQHGMDDK
ACD_SE_SE8   TPGPGTRYPL TFGWCFKLVP VD.PREVEEA NEGENNCLLH PMSQHGIEDE
ACG_BE_VI1   TPGPGTRFPL TFGWCFKLVP MD.PTEVEEA NEGENNCLLH PISQHGMEDE
AD_SE_SE69   TPGPGIRYPL IFGWCFELVP VD.PKEVEED IG.ENSSLLH PMHQHGMEDT
AD_SE_SE71   TPGPGTRFPL TFGWCFKLVP VD.PDEVEKA TEGETNTLLH PICQHGMDDE
ADHK_NO_97   TPGPGERFPL TFGWCFKLVP VD.PQEVEKA NDGENNCLLH PMCQHGMEDE
ADK_CD_MAL   TPGPGIRFPL TFGWCFKLVP MS.PEEVEEA NEGENNCLLH PISQHGMEDA
AG_BE_VI11   IPGPGTRFPL TFGWCFKLVP MD.PAGIEKA NEGENNSLLH PICQHGMGDT
AG_NG_92NG   TPGPGTRFPL TFRWCFKLVP MD.PAEIEEA NKGENNSLLH PICQHGLEDA
AGHU_GA_VI   TPGPGIRYPL CFGWCYKLVP VD.PKEVEEA TEGENNCLLH PICQQGMDDG
AGU_CD_Z32   TPGPGTRYPL CFGWCFKLVP VD.PREVEEA NTGENNCLLH PMSQHGMDDD
AJ_BW_BW21   TPGPGTRFPL TFGWCYKLVP VD.PKEVEEA NEGENNCLLH PLCQHGMEDE
   B_AU_VH   TPGPGTRYPL TFGWCFKLVP VE.PDQVEKA NEGENISLLH PMSLHGMEDK
 B_CN_RL42   TPGPGTRFPL TFGWCFKLVP VD.PEQVEEA NEGENNCLLH PMSQHGADDP
  B_DE_D31   TPGPGTRFPL TFGWCFKLVP VK.PEQVEEA NEGENNSLLH PMSLHGMDDP
  B_DE_HAN   TPGPGVRYPL TFGWCFKLVP VE.PDEEENS ......SLLH PASLHGTEDT
 B_FR_HXB2   TPGPGVRYPL TFGWCYKLVP VE.PDKIEEA NKGENTSLLH PVSLHGMDDP
  B_GA_OYI   TPGPGIRYPL CFGWCFKLVP MD.PDQVEEA NEGENNSLLH PISLHGMDDP
 B_GB_CAM1   TPGPGIRYPL TFGWCFKLVP VE.PEQVEEA NKRENASLLH PMSQHGMDDP
  B_GB_GB8   TPGPGTRFPL TFGWCFKLVP VE.PEEVEKA NEGENNCLLH PMSQHGIEDP
 B_GB_MANC   TRGPGIRYPL AFGWCFKLVP VD.PEQVEEA NEGENNSLLH PMSLHGMDDP
   B_KR_WK   TPGPGTRFPL TFGWCFKLVP VE.PEKVEEA TVGKNNCLLH PMNLHGMDDP
B_NL_3202A   TPGPGIRYPL TFGWCFKLVP VE.QEKIEEA NEGENNSLLH PMSQHGMDDP
B_TW_TWCYS   TPGPGVRYPL TFGWCFKLVP VE.PEQVEKA NEGENXCLLH PMSQHGMDDP
   B_US_BC   TPGPGIRYPL TFGWCFKLVP VD.PEKIEEA NEGENNSLLH PMSQHGMDDP
B_US_DH123   TPGPGIRYPL TFGWCFKLVP VD.PEKVEAA NEGENNCLLH PISLHGMEDP
B_US_JRCSF   TAGPGVRFPL TFGWCFKLVP VD.PEKVEEA NEGENNCLLH PMSQHGMDDP
 B_US_MNCG   TPGPGIRYPL TFGWCFKLVP VE.PEKIEEA NKGENNCLLH PMSQH.MDDP
 B_US_P896   TPGPGIRYPL TFGWCYKLVP VE.PDEGENN RE..DNSLLH PANQHGVEDS
   B_US_RF   TPGPGTRYPL TFGWCFKLVP VE.PDKVEEA TEGENNSLLH PICLHGMDDP
  B_US_SF2   TPGPGIRYPL TFGWCFKLVP VE.PEKVEEA NEGENNSLLH PMSLHGMEDA
B_US_WEAU1   TPGPGTRYPL CFGWCFKLVP VE.PEKVEEA NEGENNSLLH PMSLHGMDDH
 B_US_WR27   TPGPGTRYPL TFGWCFKLVP LE.PDQVEEA NKGENNCLLH PMSQHGMDDP
  B_US_YU2   TPG.GTRWPL TFGWCFKLVP VE.PEKIEEA NAGENNCLLH PMSQHGMDDP
BF1_BR_93B   TPGPGTRYPL TLGWCFKLVP VD.PEEVEKA NEGENNCLLH PMSQHGMEDE
C_BR_92BR0   TPGPGVRFPL TFGWCFKLVP VD.PREVEEA NTGENNSLLH PMSLHGMEDS
C_BW_96BW0   TPGPGVRYPL TFGWCYKLVP VD.PKEVEEA NEGENNRLLH PMSLHGMEDA
C_BW_96BW1   TPGPGVRYPL TFGWCFKLVP VD.PGEVEEA NEGEN.CLLH PIAQHGMEDE
C_BW_96BW1   TPGPGVRYPL TFGWCFKLVP VD.PGEVEEA NGGEDNCLLH PMSQHGIEDA
C_BW_96BW1   TPGPGVRYPL TFGWCYKLVP VD.PREVEEN NQGENHCLLH PMSQHGMDNP
C_ET_ETH22   TPGPGVRYPL TFGWCFKLVP VD.PSEVEEI NEGENNCLLH PASLHGMEDE
C_IN_93IN1   TPGPGVRYPL TFGWCFKLVP VD.PREVEEA NEGEDNCLLH PVCQHGMEDE
C_IN_93IN9   TPGPGVRYPL TFGWCFKLVP VD.PREVEEA NEGEDNCLLH PVCQHGMDDE
C_IN_93IN9   TPGPGVRFPL TFGWCFKLVP VD.PREVEEA NEGENNCLLH PVCQHGMEDE
C_IN_94IN1   TPGPGTRFPL TFGWCFKLVP VD.PREVEEA NEGEDNCLLH PVCQHGMEDD
C_IN_95IN2   TPGPGVRFPL TFGWCFKLVP VD.PREVEEA NKGEDNCLLH PVCQHGMEDD
```

```
CRF01_AE_C   TPGPGIRYPL CFGWCYKLVP VD.PKEVEED NKDESNCLLH PMSQHGVDDE
CRF01_AE_C   TPGPGVRYPL CFGWCFKLVP VD.PREVEED NKGENNCLLH PMSQHGIDDD
CRF01_AE_C   TPGPGTRYPL CFGWCYKLVP VD.PREVEED NKGENNCLLH PMSQHGIEDE
CRF01_AE_T   TPGPGIRYPL CFGWCFKLVP VD.PREVEED NKGENNCLLH PLSQHGIEDE
CRF01_AE_T   TPGPGIRYPL CFGWCFKLVP VD.PREVEED NKEENNCLLH PMSQHGIDDE
CRF01_AE_T   TPGPGIRYPL CFGWCFKLVP VD.PREVEED NKGENNCLLH PMSQHGIEDE
CRF01_AE_T   TPGPGIRFPL CFGWCFKLVP VD.PREVEED NKGENNCLLH PMSQHGMGDG
CRF01_AE_T   TPGPGIRFPL CFGWCFKLVP VD.QREVEED NKGENNCLLH PMSQHGIEDE
CRF01_AE_T   TPGPGVRLPL CFGWCFKLVP VD.PREVEED NKGENNSLLH PMSQHGIEDE
CRF02_AG_F   TPGPGTRYPL TFGWCFNLEP ID.PAEIEEA NKEENNSLLH PMSQHGIEDE
CRF02_AG_F   TPGPGTRYPL TFGWCFKLEP MD.PAEVEEA NKGENNSLLH PICQHGMEDE
CRF02_AG_G   TPGPGTRFPL TFGWCFKLVP MD.PKAVEEA TEGENNSLLH PICQHGIEDE
CRF02_AG_N   TPGPGTRFPL TFGWCFKLVP MD.PAEVEEA NEGENNSLLH PICQHGMEDD
CRF02_AG_S   TPGPGTRFPL TFGWCFKLVP MD.PADIEKD TEGENNSLLH PICQHGMEDE
CRF02_AG_S   TPGPGIRYPL TFGWCYKLVP MD.PAEVEEA NQGENNSLLH PICQHGMEDE
CRF03_AB_R   ...PGIRFPL TFGWCYKLVP VD.PAEVEEA TEGENNSLLH PICQHGMDDE
CRF03_AB_R   TPGPGIRFPL TFGWCYKLVP VD.PDEVEEA TEGENNSLLH PICQHGMDDE
CRF04_cpx_   TPGPGERFPL CFGWCFKLVP VD.PQEVEEA TEGENTCLLH PISQHGMEDE
CRF04_cpx_   TPGPGERFPL CFGWCFKLVP VD.PQEVEEA NEGENNCLLH PISQHGMEDE
CRF04_cpx_   TPGPGTRFPL CFGWCFKLVP VD.PQEVEEI TAGEDNCLLH PISQHGMEDE
CRF05_DF_B   TPGPGIRYPL TLGWCFKLVP VN.PEEVEKA NEGEDNCLLH PMSLHGMEDD
CRF05_DF_B   TPGPGTRYPL TFGWCFKLVP VD.PEEVEKA NEGENKCLLH PMHQHGMDDE
CRF06_cpx_   TPGPGTRFPL TFGWCYKLVP VD.PEEVEEL TKGENNCLLH PICQHGAEDE
CRF06_cpx_   TPGPGIRYPL TFGWCYKLVP VD.PREVEEE TKGENNSLLH PMCQHGAEDE
CRF06_cpx_   TPGPGIRYPL TFGWCYKLVP VD.PREVEED TKGENNCLLH PMSQHGMEDE
CRF06_cpx_   TPGPGIRYPL TFGWCFKLVP VD.PKEVEEE TKGENXSLLH PMCQHGVDDP
CRF11_cpx_   TPGPGVRYPL CFGWCYKLVP VD.PREVEEA NEGENNCLLH PMSQHGMDDE
CRF11_cpx_   TPGPGIRYPL CFGWCYKLVP VD.PREVEEA NEGENNCLLH PMSQHGIEDE
D_CD_84ZR0   TPGPGIRYPL TFGWCFELVP VD.PEVVEKA TEGEDNCLLH PICQHGMEDE
  D_CD_ELI   TPGPGIRYPL TFGWCYELVP VD.PQEVEED TEGETNSLLH PICQHGMEDP
  D_CD_NDK   TPGPGIRYPL TFGWCFQLVP VD.PQEVEEA TEREDNCLLH PMCQQGMEDP
D_UG_94UG1   TPGPGIRYPL TFGWCFELVP ME.PKEVEEN TEGEDNCLLH PINQHGMEDP
F1_BE_VI85   TPGPGIRYPL TLGWCFKLVP VD.PEEVEKA NEGENNCLLH PMSQHGMEDE
F1_BR_93BR   TPGPGIRYPL TMGWCFKLVP VD.PEEVEKA NEGENNCLLH PMSQHGMEDE
F1_FI_FIN9   TPGPGVRYPL TFGWCFKLVP VE.PEEVEKA NEGENNCLLH PMSQHGMEDE
F1_FR_MP41   TPGPGIRFPL TFGWCFKLVP VD.PDEVEKA NEGENNCLLH PMSQHGMDDE
F2_CM_MP25   TPGPGPRFPL TFGWCFKLVP VD.PEEVEKA NEGENNCLLH PMSLHGMEDD
F2KU_BE_VI   TPGPGIRYPL CFGWCFKLVP MD.PQEVEEA NVGENNSLLH PICQHGIDDT
G_BE_DRCBL   TPGPGTRVPL TFGWCFKLVP ME.PSEVEEA NKGENNSLLH PICQHGMEDE
G_NG_92NG0   TPGPGTRLPL TFGWCFKLVP MD.PAEIEEA NKGENISLLH PICQHGMEDE
G_SE_SE616   TPGPGTRFPL TFGWCFKLVP MD.PAEVEEA NKGENNSLLH PICQHGMEDE
H_BE_VI991   TPGPGERYPL TFGWCFKLVP VD.PQDVEKA NEGENNSLLH PMCQHGIEDP
H_BE_VI997   TPGPGEGYPL TFGWCFKLIP VD.PQEVERA NEGENNCLLY PICQHGMEDE
H_CF_90CF0   TPGPGERFPL TFGWCFKLVP VN.PQEVEQA NEGENNSLLH PMSLHGMEDD
J_SE_SE702   TPGPGTXYPL TFGWCFKLVP VD.PSEVEEA NEGENNCLLH PACQHGIEDE
J_SE_SE788   TPGPGIRYPL TFGWCYKLVP VD.PSEVEEA NEGENNCLLH PICQHGIEDE
K_CD_EQTB1   TPGPGIRYPL TFGWCYKLVP VD.PREVEEA TEGENNCLLH PVNQHGMEDE
K_CM_MP535   TPGPGIRYPL TFGWCYKLVP VD.PAEVEET TEGEDNCLLH PINQHGMEDE
N_CM_YBF30   TPGPGIRYPV TFGWCFKLVP LS.AEEVEEA NEGDNNALLH PICQHGADDD
O_CM_ANT70   TPGPGTRFPL TFGWLFKLVP VSEEEAERLG NTCERANLLH PACAHGFEDT
O_CM_MVP51   TPGPGPRFPL TFGWLFKLVP VSAEEAERLG NTNEDASLLH PACNHGAEDA
O_SN_MP129   TPGPGPRFPL TFGWLFKLVP VSEAEAEELG NKCERASLLH PACNHGFEDN
O_SN_MP130   TPGPGTRFPL TFGWLFKLVP VSEAEAEELG NKCDRAKLLH PVCNHGFEDP
U_CD___83C   TPGPGIRYPL TFGWPFKLVP VD.PKEVEEA NEGENNCLLH PICQHGMDDE

             201                                          232
00BW0762_1   HREVLMWKFD SSLARRHVAR E.LHPEYYKD C.
00BW0768_2   EREVLRWKFD SSLARRHMAR E.LHPEYYKD C.
00BW0874_2   DREVLKWQFD SSLVRRHMAR E.LHPEYYKD C.
00BW1471_2   DREVLKWVFD SSLARRHMAR E.LHPEYYKD C.
00BW1616_2   DREVLRWKFD SSLARRHLAR E.LHPEYYKD C.
```

```
00BW1686_8  DREVLMWKFD SHLAYRHMAR E.LHPEYYKD C.
00BW1759_3  HGEVLMWKFD SALARRHMAR E.LHPEYYKD C.
00BW1773_2  HGEVLKWKFD SSLARRHLAR E.KHPEFYKD C.
00BW1783_5  DKEVLQWKFD SSLARRHMAR E.LHPEYYKD C.
00BW1795_6  DREVLMWKFD SHLARRHTAR E.LHPEFYKD CR
00BW1811_3  HGEVLEWKFD SMLARRHMAR E.LHPEYYKD C.
00BW1859_5  EREVLRWKFD SQLARRHMAR E.LHPEYYKD C.
00BW1880_2  HKEVLRWKFD SSLARRHLAR E.LHPEYYKD C.
00BW1921_1  DREVLMWKFD SQLAHRHLAR E.KHPEWYKD C.
00BW2036_1  HKEVLTWKFD SHLARRHMAR E.LHPEYYKD C.
00BW2063_6  HREVLQWKFD SQLARRHIAR E.RHPEYYKD C.
00BW2087_2  DREVLKWVFD SSLARRHLAR E.KHPEFYKD C.
00BW2127_2  EREVLQWKFD SLLSRRHLAR IXIHAEYYKD C.
00BW2128_3  HGEVLMWKFD SHLAYRHMAR E.KHPEFYKD C.
00BW2276_7  EREVLKWKFD SSLARRHMAR E.LHPEYYKA C.
00BW3819_3  EGEVLQWKFD SLLAYRHMAR E.QHPEYYKD C.
00BW3842_8  DREVPMWKFD SLLAHRHMAR E.LHPGYYKD C.
00BW3871_3  DREVLMWKFD SHLVHRHMAR E.LHPEYYKG C.
00BW3876_9  DREVLRWKFD S.LARRHIAR E.LHPEYYKN C.
00BW3886_8  EREVLKWKFD SQLAYRHMAR E.IHPEYYKD C.
00BW3891_6  DREVLRWKFD SHLARRHMAR E.LHPEWYKD C.
00BW3970_2  DREVLKWQFD ISLARRHMAR E.LHPEWYKD ..
00BW5031_1  DREVLRWKFD SELARRHIAR E.RHPEFYKD C.
 96BW01B21  HREVLKWKFD SQLARRHMAR E.LHPEYYKD C.
  96BW0407  DREVLRWKFD SSLAHRHMAR E.LHPEYYKD C.
  96BW0502  HGEVLKWKFD SQLARRHMAR E.LYPEYYKD C.
 96BW06_J4  EREVLTWKFD SHLVHRPMAR E.IHPEYYKD C.
 96BW11_06  HKEVLKWKFD SQLARRHLAR E.LHPEFYKD C.
  96BW1210  DREVLKWKFD SSLARRHLTR E.KHPEYYKD C.
 96BW15B03  DKEVLMWKFD SHLARRHMAR E.LHPEYYKD C.
 96BW16_26  ERGVLKWKFD SHLARRHMAR E.LHPEYYKD C.
 96BW17A09  DREVLKWVFD SHLARKHMAR E.LHPEYYKN C.
 96BWMO1_5  HREVLKWKFD SSLARRHMAR E.LHPEFYKD C.
 96BWMO3_2  HGEVLMWKFD SQLARRHMAR E.LHPEYYKD C.
98BWMC12_2  EREVLKWQFD SSLARRHMAR E.LHPEYYKD C.
98BWMC13_4  GKEVLIWKFD SHLARRHMAR E.LHPEFYKD C.
98BWMC14_a  DREVLMWKFD SQLARRHIAR E.IHPEYYKD C.
98BWMO14_1  EKEVLKWVFD SSLARRHVAR E.LHPEFYKD C.
98BWMO18_d  DREVLRWKFD SSLARRHMAR E.LHPEYYKD C.
98BWMO36_a  DREVLKWEFD IRLAHTHMAR Q.LHPEFYKN C.
98BWMO37_d  DREVLKWQFD SSLAHRHVAR E.LHPEYYKD C.
99BW3932_1  DREVLKWKFD SMLARRHMAR E.LHPEFYKD C.
99BW4642_4  DREVLIWKYD SQLARRHMAR E.LHPDYYKD C.
99BW4745_8  DREVLRWKFD SHLARRHMAR E.LHPEFYKD C.
99BW4754_7  HKEVLKWKFD SHLARRHMAR E.LHPEFYKD C.
99BWMC16_8  DREVLKWQFD SSLARRHMAR E.LHPEYYKD C.
A2_CD_97CD  EREVLKWKFD SRLALRHLAR E.QHPEFYKD C.
A2_CY_94CY  EREVLRWEFD RSLARRHRAR E.LHPEYYKD C.
A2D___97KR  EREVLKWVFD SHLALVHKAR E.LHPEFYKD C.
A2G_CD_97C  DKQVLGWRFD SSLARRHIAR E.KHPEYYKD C.
A_BY_97BL0  EKEVLMWKFD SRLALKHRAR E.LHPEFYKD C.
  A_KE_Q23  EREVLKWKFD SRLALKHRAR E.LHPEWYKD C.
A_SE_SE659  EKEVLKWKFD SRLALKHLAC E.KHPEFYKD C.
A_SE_SE725  EKETLRWRFD SRLALRHRAQ E.MHPEFYKD C.
A_SE_SE753  EREVLKWKFD SRLALKHRAQ E.LHPEFYKD C.
A_SE_SE853  ERETLMWKFD SKLALKHRAH E.LHPEYFKN C.
A_SE_SE889  ERETLMWKFD SRLALTHRAR E.LHPEFYKD C.
A_SE_UGSE8  ERETLMWKFD PHLAFKHRAF E.LHPEYYKN ..
A_UG_92UG0  EKETLRWKFD SSLARVHKAR E.LHPEFYKD C.
 A_UG_U455  EKEVLMWKFD STLALKHRAY E.LHPEFYKD ..
AC_IN_2130  YGEVLQWKFD SHLAYKHQAR E.RHPEFYKD C.
```

```
AC_RW_92RW   DREVLKWKFD SHLAHRHMAR E.LHPEYYKD C.
AC_SE_SE94   ERETLVWRFD SRLALKHLAR E.KHPEFYKD C.
ACD_SE_SE8   DKEVLRWKFD SQLARRHMAR E.MHPEYYKD C.
ACG_BE_VI1   DREVLVWRFD SRLALKHIAK E.KHPEYFKD C.
AD_SE_SE69   EREVLMWRFN SRLAFEHKAH Q.LHPEYYKD C.
AD_SE_SE71   EKEVLKWQFD SRLALKHLAR E.KHPEFYKD C.
ADHK_NO_97   EXEVLMWRFD SRLAFKHRAR E.LHPEFYKD C.
ADK_CD_MAL   EREVLKWKFD SSLALRHRAR E.QHPEYYKD C.
AG_BE_VI11   EREVLVWKFD SMLAFKHRAR E.LHPEYYKD C.
AG_NG_92NG   DREVLVWRFD SSLARRHIAR E.QHPEYYKD C.
AGHU_GA_VI   EREVLMWKFD SSLAREHVAR K.LYPEFFKD C.
AGU_CD_Z32   EREVLMWKFD SSLARKHLAR E.MHPEFYKD ..
AJ_BW_BW21   DREVLMWKFD SSLARRHLAR E.KHPEFYKD C.
   B_AU_VH   EKEVLMWKFD SRLAVHHMAR E.LHPEYYKN ..
 B_CN_RL42   EREVLMWKFD SRLAIHHMAR E.MHPEYHKD C.
  B_DE_D31   EREVLVWRFD SRLAFKHMAR E.LHPEYYKN ..
  B_DE_HAN   EREVLKWKFD SHLAFHHKAR E.LHPEYYKD C.
 B_FR_HXB2   EREVLEWRFD SRLAFHHVAR E.LHPEYFKN C.
  B_GA_OYI   EKEVLVWKFD SRLAFRHMAR E.VHPEYYKD C.
 B_GB_CAM1   EKEVLMWKFD SRLAFHHMAR E.KHPEFYKD C.
  B_GB_GB8   EKEVLVWKFN SRLAFHHMAR E.LHPEFYKD C.
 B_GB_MANC   EKEVLVWKFD SRLAFHHVPD E.LHPEYYKD C.
   B_KR_WK   EGEVLVWRFD SRLAFHHMAR E.KHPEYYKD C.
B_NL_3202A   EREVLEWRFD SRLAFHHMAR E.KHPEYYKD C.
B_TW_TWCYS   EKEVLVWRFD STLAFHHRAR E.LHPEYYKX C.
   B_US_BC   EREVLEWRFD SRLAFHHMAR E.LHPEYYKN R.
B_US_DH123   EKEVLLWKFD SRLAYHHMAR E.LHPEYYKN C.
B_US_JRCSF   EKEVLVWKFD SKLALHHVAR E.LHPEYYKD C.
 B_US_MNCG   EREVLVWKSD SHLAFQHYAR E.LHPEYYKN C.
 B_US_P896   ERQVLVWRFD SRLAFHHVAR E.LHPEYFKN ..
   B_US_RF   EKEVLVWKFD SRLAFHHVAR E.KHPEYYKD C.
  B_US_SF2   EKEVLVWRFD SKLAFHHMAR E.LHPEYYKD C.
B_US_WEAU1   EKEVLMWKFD SKLAFHHVAR E.LHPEYFKD C.
 B_US_WR27   EKEVLVWKFD SRLAFHHKAR E.LHPEYYKN ..
  B_US_YU2   EREGLEWRFD SRLAFHHVAR E.LHPEYYKN ..
BF1_BR_93B   DREILQWRFD SRLAFHHMAR E.LHPEYYKD C.
C_BR_92BR0   HREVLQWKFD SLLARRHMAR E.LHPEYYKD C.
C_BW_96BW0   DGEVLRWKFD SHLAHRHMAR E.LHPEYYKD C.
C_BW_96BW1   HKEVLKWKFD SQLARRHLAR E.LHPEFYKD C.
C_BW_96BW1   DREVLKWKFD SSLARRHLTR E.KHPEYYKD C.
C_BW_96BW1   DKEVLMWKFD SHLARRHMAR E.LHPEYYKD C.
C_ET_ETH22   DREVLKWKFD SHLARRHMAR E.LHPEYYKD C.
C_IN_93IN1   HREVLKWKFD SQLARRHMAR E.LHPEFYKD C.
C_IN_93IN9   HREVLQWKFD SLLAHRHRAR E.LHPEFYKD C.
C_IN_93IN9   HREVLQWKFD SHLAHRHMAR E.LHPEYYKD C.
C_IN_94IN1   HREVLMWK.. .QLAHRHIAR E.LHPEFYKD C.
C_IN_95IN2   HNEVLVWKFD SQLAHKHRAR E.LHPEFYNK DC
CRF01_AE_C   EREVLMWKFD SSLARRHIAR E.LRPEYYKD C.
CRF01_AE_C   EREVLMWKFD SSLARRHIAR E.LHPEYYKD ..
CRF01_AE_C   EREVLMWKFD SSLARRHIAR E.LHPEYYKD C.
CRF01_AE_T   EREVLMWKFD SALARKHTAR E.LHPEYYKD C.
CRF01_AE_T   EREVLMWKFD STLARKHIAR E.QHPEFYKD C.
CRF01_AE_T   EREVLIWKFD SALARRHIAR E.LRPEFYKD C.
CRF01_AE_T   EREVLMWKFD SALARKHIAR E.MHPEYYKD C.
CRF01_AE_T   EREVLMWKFD SALARKHVAR E.QHPEYYKD C.
CRF01_AE_T   EREVLIWKFD SSLARKHLAR E.LHPEYYKD C.
CRF02_AG_F   DREVLVWRFD SSLARTHRAR E.LHPEYYKD C.
CRF02_AG_F   DREVLVWRFD SSLARRHIAR E.RHPEFYKD C.
CRF02_AG_G   DREVLVWRFD SSLAFTHRAR E.MHPEFYKD C.
CRF02_AG_N   DREVLIWRFD SRLAFRHTAR E.LHPEYYKD C.
CRF02_AG_S   DREVLVWRFD SRLAFTHKAR E.MHPEFYKD CX
```

```
CRF02_AG_S  DKEVLVWRFD SRLAFRHTAR E.LHPEYYKD C.
CRF03_AB_R  EKEVLMWKFD SRLALTHRAR E.LHPEFYKD C.
CRF03_AB_R  EKEVLMWKFD SRLALTHRAR E.LHPEFYKD C.
CRF04_cpx_  EREVLKWKFD SRLAYKHVAR E.LHPEFYKD C.
CRF04_cpx_  EREVLKWKFD SRLAFKHIAR E.LHPEFYKD C.
CRF04_cpx_  EREVLKWKFD SLLAYRHMAR E.LHPEFYKD C.
CRF05_DF_B  DREVLQWKFD SSLALRHIAR E.RHPEFYQD ..
CRF05_DF_B  DGEVLRWKFD SSLALKHIAR E.RRPEFYQD ..
CRF06_cpx_  EREVLKWKFD SSLARRHIAR E.KHPEFYKD C.
CRF06_cpx_  EGEVLMWKFD SSLARRHIAR E.LHPDFYKD C.
CRF06_cpx_  EREVLMWKFD SSLARRHTAR E.MHPEFYKD C.
CRF06_cpx_  EXEVLMWKFD SSLARRHIAX E.XHPEFXKD C.
CRF11_cpx_  EREVLKWVFD SSLARKHIAR E.LHPDFYKD ..
CRF11_cpx_  DREVLRWKFD SSLARRHIAR E.LHPDFYKD ..
D_CD_84ZR0  EKEVLVWRFN SRLAFEHKAK E.KYPEYFKN C.
  D_CD_ELI  ERQVLKWRFN SRLAFEHKAR E.MHPEFYKN ..
  D_CD_NDK  ERQVLMWRFN SRLALEHKAR E.LHPEFYKD C.
D_UG_94UG1  EREVLVWRFN SRLAFEHKAK M.KHPEYYKD C.
F1_BE_VI85  DREVLRWKFD SSLALRHIAR E.RHPEFYQD ..
F1_BR_93BR  DKEVLKWEFD SRLALRHIAR E.RHPEYYQD ..
F1_FI_FIN9  DREVLKWKFD SRLALKHIAR E.RHPEFYRD ..
F1_FR_MP41  DREVLRWEFD SRLAFRHIAR E.KHPEFYQN ..
F2_CM_MP25  DKEVLKWQFD SRLALRHIAR E.RHPEYYKD ..
F2KU_BE_VI  EREVLVWKFD SRLALKHLAR E.KHPEYYKD C.
G_BE_DRCBL  DGEVLVWRFD SSLARRHLAR E.LHPEYYKD C.
G_NG_92NG0  DREVLVWRFN SSLARRHLAR E.LHPEYYKD C.
G_SE_SE616  DREVLVWRFD SSLARRHIAR E.LHPEYYKD C.
H_BE_VI991  EREVLMWKFD SRLALRHRAK E.LHPEFYKD C.
H_BE_VI997  EGEVLMWKFD SRLAFTHTAR E.KHPEFYKD C.
H_CF_90CF0  GREVLMWKFD SRLALTHLAR V.KHPEY.KD C.
J_SE_SE702  EREVLKWKFD SSLARRHIAR E.LHPEFYKD C.
J_SE_SE788  EREVLQWKFD SSLARRHIAR E.LHPEFYKD C.
K_CD_EQTB1  HREVLKWKFD SSLARKHVAR E.MHPEYYKD ..
K_CM_MP535  HREILMWKFD SSLARRHVAR E.LHPDYYKD ..
N_CM_YBF30  HKEVLVWRFD SSLARRHVAR E.LHPEFYKN C.
O_CM_ANT70  HKEILMWKFD RSLGNTHVAM ITHPELFQKD ..
O_CM_MVP51  HGEILKWQFD RSLGLTHIAL QKHPELFPSN ..
O_SN_MP129  HGQILKWQFD RSLGSTHVAM VTNPELFNKD ..
O_SN_MP130  HKEMLKWQFD RSLGSTHVAL ITHPELFLKD ..
U_CD___83C  EKEVLMWKFD SSLARRHLAR E.LHPEFYKD C.
```

Table 14. HIV Pol Sequence Alignment

GCG Multiple Sequence File.
Written by Omiga 1.1

Name: 00BW0762_1	Len:	1046	Check:	4376	Weight:	1.00
Name: 00BW0768_2	Len:	1046	Check:	8430	Weight:	1.00
Name: 00BW0874_2	Len:	1046	Check:	8925	Weight:	1.00
Name: 00BW1471_2	Len:	1046	Check:	1324	Weight:	1.00
Name: 00BW1616_2	Len:	1046	Check:	935	Weight:	1.00
Name: 00BW1686_8	Len:	1046	Check:	8131	Weight:	1.00
Name: 00BW1759_3	Len:	1046	Check:	579	Weight:	1.00
Name: 00BW1773_2	Len:	1046	Check:	1975	Weight:	1.00
Name: 00BW1783_5	Len:	1046	Check:	216	Weight:	1.00
Name: 00BW1795_6	Len:	1046	Check:	5932	Weight:	1.00
Name: 00BW1811_3	Len:	1046	Check:	6525	Weight:	1.00
Name: 00BW1859_5	Len:	1046	Check:	2879	Weight:	1.00
Name: 00BW1880_2	Len:	1046	Check:	7093	Weight:	1.00
Name: 00BW1921_1	Len:	1046	Check:	2524	Weight:	1.00
Name: 00BW2036_1	Len:	1046	Check:	8279	Weight:	1.00
Name: 00BW2063_6	Len:	1046	Check:	3935	Weight:	1.00
Name: 00BW2087_2	Len:	1046	Check:	7898	Weight:	1.00
Name: 00BW2127_2	Len:	1046	Check:	728	Weight:	1.00
Name: 00BW2128_3	Len:	1046	Check:	5356	Weight:	1.00
Name: 00BW2276_7	Len:	1046	Check:	9456	Weight:	1.00
Name: 00BW3819_3	Len:	1046	Check:	6369	Weight:	1.00
Name: 00BW3842_8	Len:	1046	Check:	4573	Weight:	1.00
Name: 00BW3871_3	Len:	1046	Check:	6948	Weight:	1.00
Name: 00BW3876_9	Len:	1046	Check:	6609	Weight:	1.00
Name: 00BW3886_8	Len:	1046	Check:	8244	Weight:	1.00
Name: 00BW3891_6	Len:	1046	Check:	5718	Weight:	1.00
Name: 00BW3970_2	Len:	1046	Check:	3940	Weight:	1.00
Name: 00BW5031_1	Len:	1046	Check:	2442	Weight:	1.00
Name: 96BW01B21	Len:	1046	Check:	2358	Weight:	1.00
Name: 96BW0407	Len:	1046	Check:	8537	Weight:	1.00
Name: 96BW0502	Len:	1046	Check:	3948	Weight:	1.00
Name: 96BW06_J4	Len:	1046	Check:	7173	Weight:	1.00
Name: 96BW11_06	Len:	1046	Check:	973	Weight:	1.00
Name: 96BW1210	Len:	1046	Check:	5817	Weight:	1.00
Name: 96BW15B03	Len:	1046	Check:	5157	Weight:	1.00
Name: 96BW16_26	Len:	1046	Check:	3303	Weight:	1.00
Name: 96BW17A09	Len:	1046	Check:	1256	Weight:	1.00
Name: 96BWMO1_5	Len:	1046	Check:	5593	Weight:	1.00
Name: 96BWMO3_2	Len:	1046	Check:	3661	Weight:	1.00
Name: 98BWMC12_2	Len:	1046	Check:	7159	Weight:	1.00
Name: 98BWMC13_4	Len:	1046	Check:	3254	Weight:	1.00
Name: 98BWMC14_a	Len:	1046	Check:	5638	Weight:	1.00
Name: 98BWMO14_1	Len:	1046	Check:	7680	Weight:	1.00
Name: 98BWMO18_d	Len:	1046	Check:	1619	Weight:	1.00
Name: 98BWMO36_a	Len:	1046	Check:	8852	Weight:	1.00
Name: 98BWMO37_d	Len:	1046	Check:	4750	Weight:	1.00
Name: 99BW3932_1	Len:	1046	Check:	5391	Weight:	1.00
Name: 99BW4642_4	Len:	1046	Check:	1514	Weight:	1.00
Name: 99BW4745_8	Len:	1046	Check:	52	Weight:	1.00
Name: 99BW4754_7	Len:	1046	Check:	4905	Weight:	1.00
Name: 99BWMC16_8	Len:	1046	Check:	1544	Weight:	1.00
Name: A2_CD_97CD	Len:	1046	Check:	9703	Weight:	1.00
Name: A2_CY_94CY	Len:	1046	Check:	3235	Weight:	1.00
Name: A2D___97KR	Len:	1046	Check:	3776	Weight:	1.00
Name: A2G_CD_97C	Len:	1046	Check:	2059	Weight:	1.00
Name: A_BY_97BL0	Len:	1046	Check:	2724	Weight:	1.00
Name: A_KE_Q23_A	Len:	1046	Check:	1835	Weight:	1.00

Name:	A_SE_SE659	Len:	1046	Check:	647	Weight:	1.00
Name:	A_SE_SE725	Len:	1046	Check:	263	Weight:	1.00
Name:	A_SE_SE753	Len:	1046	Check:	2271	Weight:	1.00
Name:	A_SE_SE853	Len:	1046	Check:	5036	Weight:	1.00
Name:	A_SE_SE889	Len:	1046	Check:	8414	Weight:	1.00
Name:	A_SE_UGSE8	Len:	1046	Check:	3268	Weight:	1.00
Name:	A_UG_92UG0	Len:	1046	Check:	2007	Weight:	1.00
Name:	A_UG_U455_	Len:	1046	Check:	2277	Weight:	1.00
Name:	AC_IN_2130	Len:	1046	Check:	5353	Weight:	1.00
Name:	AC_RW_92RW	Len:	1046	Check:	4695	Weight:	1.00
Name:	AC_SE_SE94	Len:	1046	Check:	4206	Weight:	1.00
Name:	ACD_SE_SE8	Len:	1046	Check:	7281	Weight:	1.00
Name:	ACG_BE_VI1	Len:	1046	Check:	1400	Weight:	1.00
Name:	AD_SE_SE69	Len:	1046	Check:	4640	Weight:	1.00
Name:	AD_SE_SE71	Len:	1046	Check:	1057	Weight:	1.00
Name:	ADHK_NO_97	Len:	1046	Check:	3502	Weight:	1.00
Name:	ADK_CD_MAL	Len:	1046	Check:	2578	Weight:	1.00
Name:	AG_BE_VI11	Len:	1046	Check:	8416	Weight:	1.00
Name:	AG_NG_92NG	Len:	1046	Check:	9397	Weight:	1.00
Name:	AGHU_GA_VI	Len:	1046	Check:	9562	Weight:	1.00
Name:	AGU_CD_Z32	Len:	1046	Check:	8398	Weight:	1.00
Name:	AJ_BW_BW21	Len:	1046	Check:	3451	Weight:	1.00
Name:	B_AU_VH_AF	Len:	1046	Check:	2033	Weight:	1.00
Name:	B_CN_RL42_	Len:	1046	Check:	1369	Weight:	1.00
Name:	B_DE_D31_U	Len:	1046	Check:	4607	Weight:	1.00
Name:	B_DE_HAN_U	Len:	1046	Check:	1771	Weight:	1.00
Name:	B_FR_HXB2_	Len:	1046	Check:	4569	Weight:	1.00
Name:	B_GA_OYI__	Len:	1046	Check:	3682	Weight:	1.00
Name:	B_GB_CAM1_	Len:	1046	Check:	3161	Weight:	1.00
Name:	B_GB_GB8_A	Len:	1046	Check:	6253	Weight:	1.00
Name:	B_GB_MANC_	Len:	1046	Check:	7670	Weight:	1.00
Name:	B_KR_WK_AF	Len:	1046	Check:	8737	Weight:	1.00
Name:	B_NL_3202A	Len:	1046	Check:	2083	Weight:	1.00
Name:	B_TW_TWCYS	Len:	1046	Check:	3056	Weight:	1.00
Name:	B_US_BC_L0	Len:	1046	Check:	3160	Weight:	1.00
Name:	B_US_DH123	Len:	1046	Check:	1102	Weight:	1.00
Name:	B_US_JRCSF	Len:	1046	Check:	5571	Weight:	1.00
Name:	B_US_MNCG_	Len:	1046	Check:	3988	Weight:	1.00
Name:	B_US_P896_	Len:	1046	Check:	2465	Weight:	1.00
Name:	B_US_RF_M1	Len:	1046	Check:	3672	Weight:	1.00
Name:	B_US_SF2_K	Len:	1046	Check:	1754	Weight:	1.00
Name:	B_US_WEAU1	Len:	1046	Check:	2993	Weight:	1.00
Name:	B_US_WR27_	Len:	1046	Check:	4098	Weight:	1.00
Name:	B_US_YU2_M	Len:	1046	Check:	5564	Weight:	1.00
Name:	BF1_BR_93B	Len:	1046	Check:	4182	Weight:	1.00
Name:	C_BR_92BR0	Len:	1046	Check:	5481	Weight:	1.00
Name:	C_BW_96BW0	Len:	1046	Check:	6833	Weight:	1.00
Name:	C_BW_96BW1	Len:	1046	Check:	2166	Weight:	1.00
Name:	C_BW_96BW1	Len:	1046	Check:	5817	Weight:	1.00
Name:	C_BW_96BW1	Len:	1046	Check:	5157	Weight:	1.00
Name:	C_ET_ETH22	Len:	1046	Check:	3509	Weight:	1.00
Name:	C_IN_93IN1	Len:	1046	Check:	5471	Weight:	1.00
Name:	C_IN_93IN9	Len:	1046	Check:	4102	Weight:	1.00
Name:	C_IN_93IN9	Len:	1046	Check:	3150	Weight:	1.00
Name:	C_IN_94IN1	Len:	1046	Check:	5157	Weight:	1.00
Name:	C_IN_95IN2	Len:	1046	Check:	4641	Weight:	1.00
Name:	CRF01_AE_C	Len:	1046	Check:	87	Weight:	1.00
Name:	CRF01_AE_C	Len:	1046	Check:	3758	Weight:	1.00
Name:	CRF01_AE_C	Len:	1046	Check:	2775	Weight:	1.00
Name:	CRF01_AE_T	Len:	1046	Check:	1864	Weight:	1.00
Name:	CRF01_AE_T	Len:	1046	Check:	7414	Weight:	1.00

```
Name: CRF01_AE_T      Len:  1046  Check: 7837  Weight:  1.00
Name: CRF01_AE_T      Len:  1046  Check: 3529  Weight:  1.00
Name: CRF01_AE_T      Len:  1046  Check: 7503  Weight:  1.00
Name: CRF01_AE_T      Len:  1046  Check: 5730  Weight:  1.00
Name: CRF02_AG_F      Len:  1046  Check: 9432  Weight:  1.00
Name: CRF02_AG_F      Len:  1046  Check: 2064  Weight:  1.00
Name: CRF02_AG_G      Len:  1046  Check: 9849  Weight:  1.00
Name: CRF02_AG_N      Len:  1046  Check: 1793  Weight:  1.00
Name: CRF02_AG_S      Len:  1046  Check: 4817  Weight:  1.00
Name: CRF02_AG_S      Len:  1046  Check: 1764  Weight:  1.00
Name: CRF03_AB_R      Len:  1046  Check: 1695  Weight:  1.00
Name: CRF03_AB_R      Len:  1046  Check: 1425  Weight:  1.00
Name: CRF04_cpx_      Len:  1046  Check: 8496  Weight:  1.00
Name: CRF04_cpx_      Len:  1046  Check: 2074  Weight:  1.00
Name: CRF04_cpx_      Len:  1046  Check: 9245  Weight:  1.00
Name: CRF05_DF_B      Len:  1046  Check: 62  Weight:  1.00
Name: CRF05_DF_B      Len:  1046  Check: 3427  Weight:  1.00
Name: CRF06_cpx_      Len:  1046  Check: 142  Weight:  1.00
Name: CRF06_cpx_      Len:  1046  Check: 6688  Weight:  1.00
Name: CRF06_cpx_      Len:  1046  Check: 8524  Weight:  1.00
Name: CRF06_cpx_      Len:  1046  Check: 4725  Weight:  1.00
Name: CRF11_cpx_      Len:  1046  Check: 2194  Weight:  1.00
Name: CRF11_cpx_      Len:  1046  Check: 8466  Weight:  1.00
Name: D_CD_84ZR0      Len:  1046  Check: 515  Weight:  1.00
Name: D_CD_ELI_K      Len:  1046  Check: 2096  Weight:  1.00
Name: D_CD_NDK_M      Len:  1046  Check: 3376  Weight:  1.00
Name: D_UG_94UG1      Len:  1046  Check: 3505  Weight:  1.00
Name: F1_BE_VI85      Len:  1046  Check: 3993  Weight:  1.00
Name: F1_BR_93BR      Len:  1046  Check: 2251  Weight:  1.00
Name: F1_FI_FIN9      Len:  1046  Check: 9772  Weight:  1.00
Name: F1_FR_MP41      Len:  1046  Check: 1447  Weight:  1.00
Name: F2_CM_MP25      Len:  1046  Check: 2842  Weight:  1.00
Name: F2KU_BE_VI      Len:  1046  Check: 5026  Weight:  1.00
Name: G_BE_DRCBL      Len:  1046  Check: 5377  Weight:  1.00
Name: G_NG_92NG0      Len:  1046  Check: 6000  Weight:  1.00
Name: G_SE_SE616      Len:  1046  Check: 7901  Weight:  1.00
Name: H_BE_VI991      Len:  1046  Check: 9107  Weight:  1.00
Name: H_BE_VI997      Len:  1046  Check: 5776  Weight:  1.00
Name: H_CF_90CF0      Len:  1046  Check: 9201  Weight:  1.00
Name: J_SE_SE702      Len:  1046  Check: 9700  Weight:  1.00
Name: J_SE_SE788      Len:  1046  Check: 8817  Weight:  1.00
Name: K_CD_EQTB1      Len:  1046  Check: 3723  Weight:  1.00
Name: K_CM_MP535      Len:  1046  Check: 3729  Weight:  1.00
Name: N_CM_YBF30      Len:  1046  Check: 3336  Weight:  1.00
Name: O_CM_ANT70      Len:  1046  Check: 9461  Weight:  1.00
Name: O_CM_MVP51      Len:  1046  Check: 2986  Weight:  1.00
Name: O_SN_99SE_      Len:  1046  Check: 377  Weight:  1.00
Name: O_SN_99SE_      Len:  1046  Check: 9312  Weight:  1.00
Name: U_CD___83C      Len:  1046  Check: 1358  Weight:  1.00
//

            1                                                  50
00BW0762_1  FFRENLAFPQ G.EAREFPPE QT........ ....RANSPT SR.......E
00BW0768_2  FFRENLAFPQ .GEAGEFPSE .......... ..QTRANSTT SR.......K
00BW0874_2  FFRENLAFPQ G.EAREFPPE QA........ ....RAISPT SR.......E
00BW1471_2  FFRENLAFSE G.EARELPSE Q......... ...ARAISPT SR.......E
00BW1616_2  FFRENLAFPQ G.KAGEFPPE QTRANSP... .SSTSANSPT SR.......E
00BW1686_8  FFRENLAFPQ G.EAREFPSE Q......... ...ARAISPT SR.......E
00BW1759_3  FFRENLAFPQ .GEAREFPSE .......... ..QTRANSPT TR.......E
00BW1773_2  FFRENLAFPQ G.EAREFPSE QTRAN..... .......SPT SR.......E
```

```
00BW1783_5   FFRENLAFPE GGEAREFPAE QTSRE..... ..QTRANSPT SR.......E
00BW1795_6   FFRENLAFPQ G.EAREFPSE QT........ ....RANSPT SR.......E
00BW1811_3   FFRENLAFPQ G.EAREFPSE QARANSPTR. .....ANSPT SR.......E
00BW1859_5   FFRENLAFPQ G.KAREFSPE QA........ ....RANSPT SR.......E
00BW1880_2   FFRENLAFPQ G.EAREFPPE QT........ ....RADSPT SR.......E
00BW1921_1   FFRENLAFPQ G.EAREFPSE Q......... ...ARANSST SR.......E
00BW2036_1   FFRENLAFQQ G.KAREFPSE QNSP...... ..TRRANSPT SR.......E
00BW2063_6   FFRENLAFPQ G.EAREFPSE QT........ ....RANSPT SR.......K
00BW2087_2   FFRENLAFPQ GGEAGEFPSE .......... ..QTRANSPT SR.......A
00BW2127_2   FFGENLAFPQ G.EAREFPPE QARTNSP... ..QAGAISPT SR.......E
00BW2128_3   FFRENLAFQQ .GEAREFPSE QTRTNSPTSR .EQTRANSPT SG.......E
00BW2276_7   FFRETLAFQQ G.KARELPSE QDRANSPTR. .....ANSPT GR.......Q
00BW3819_3   FFRENLAFPQ G.EAREFPPK QARTNSP... ......NSPT SR.......E
00BW3842_8   FFREDLAFPR R.KAREFPSE QNRAN..... .SPTRANSPT SR.......E
00BW3871_3   FFRENLAFPQ G.EAREFPSE Q......... ...TRANSPT SR.......K
00BW3876_9   FFRENLAFPQ G.KAREFPSK QA........ ....RANSPT GR.......E
00BW3886_8   FFRENLAFPQ G.EAREFPSE QTRANSPT.. ...SRANSPT SR.......E
00BW3891_6   FFRENLAFPQ G.EAREFSSE .....Q.... ...ARANSPT SR.......E
00BW3970_2   FFREILAFPE G.EAWEFPSE Q......... ...IRANSPT SR.......E
00BW5031_1   FFRENLAFQQ G.EARELPPE Q...TRTNS. ..PTNANSPT SR.......E
 96BW01B21   FFRENLAFPQ G.KAREFPSE Q......TR. .....AISPT SR.......E
  96BW0407   FFRENLAFPQ G.EAREFPSE Q......... ...TRANSPT SR.......E
  96BW0502   FFRENLAFPQ G.EAREFPPE QIRASSPNS. .....TNSPT SR.......E
 96BW06_J4   FFRENLAFPQ RGEAREFPSE .......... ..QARANSPT SR.......E
 96BW11_06   FFRENLAFPQ G.EAREFPSE .......... ..QTGANSPT SR.......E
  96BW1210   FFRENLAFPQ G.EAREFPSE QTRAIS.... ........PT SR.......E
 96BW15B03   FFREDLAFPQ G.KAREFPSE QN........ ....RANSPT SR.......E
 96BW16_26   FFRENLAFPQ .GEAREFPSE .......... ..QTRANSPT SG.......E
 96BW17A09   FFRENLAFPQ GGEAREFPSE Q......... ...ARANSPT SR.......E
 96BWMO1_5   FFRENLAFPQ G.EAREFPSE QT........ ....RANSPT SR.......N
 96BWMO3_2   FFRENLAFPQ G.EAREFPPE QT........ ....RANSPT SR.......A
98BWMC12_2   FFRETLAFPQ G.EAREFSSE QG........ ....RANSPT SR.......E
98BWMC13_4   FFRENLAFPQ G.EAREFPSE QT........ ....RANSPT SR.......K
98BWMC14_a   FFRENLAFPQ G.EARELPSE Q......... ...TRTISPT SR.......E
98BWMO14_1   FFRENLAFPQ RGEAGEFPSE .......... ..KTRANSPT SR.......E
98BWMO18_d   FFRENLAFPQ G.EAGKFHSE QTSANSP... ..TSRANSPT SR.......E
98BWMO36_a   FFRENLAFPQ G.EAREFPPE QTRANSP... ..TSRANSPT GR.......E
98BWMO37_d   FFRENLAFPQ G.EAREFPSE K......... ...TRANSPT GR.......E
99BW3932_1   FFRENLAFQQ G.EAREFPPE QDSANSPTSR ELQDRANSPT SR.......E
99BW4642_4   FFRENLAFPQ G.EAREFLPE QD........ ....RANSPT SR.......E
99BW4745_8   FFRENLAFQQ G.EAREFPSE QTRANSP... ...TRANSPT SR.......E
99BW4754_7   FFRKNLAFQQ G.EAREFPSE QT........ ....RANSPT SR.......E
99BWMC16_8   FFREDLAFQQ R.EAREFPSE Q.TRANS... ..PTRANSPT SR.......E
A2_CD_97CD   FFRENLAFQQ R.EAREFSSE .......... ..QDRANSPT N.........
A2_CY_94CY   FFRENLAFQQ R.EARKFSSE .......... ..QNRANSPT SR.......E
A2D___97KR   FFRENLAFPQ R.EAREFSSE .......... ..QNRTNSPT SR.......G
A2G_CD_97C   FFRENLAFQQ R.EAREFS.. .......... SEQDRANSPT RR.......E
A_BY_97BL0   FFRKNLAFQQ R.EARKFSSE .......... ..QTRAISPT S.......RK
A_KE_Q23_A   FFRENLAFQK G.EAREFSSE .......... ..QTGTNSST S.......RD
A_SE_SE659   FFRENLAFQQ R.EARKFSSE .......... ..QTRANSPT S.......RD
A_SE_SE725   FFRENVAFQQ G.EARKFSSE .......... ..QTGANSPT S.......RA
A_SE_SE753   FFRENLAFQQ G.EAGKFSSE .......... ..QTGANSPT S.......RD
A_SE_SE853   FFRENLAFQQ R.EARKFSSE .......... ..QTRANSPT S.......RD
A_SE_SE889   FFRENLAFQQ G.EARKFSSE .......... ..QTGANSPT S.......RD
A_SE_UGSE8   FFRENLAFPQ G.EAGKFSSE .......... ..QTGAISPT S.......RD
A_UG_92UG0   FFRENLAFQQ R.EARKFSSE .......... ..QTRTNSPT SS......RD
A_UG_U455_   FFRENLAFQQ G.EAREFSSE .......... ..QTRANSPT SR......N.
AC_IN_2130   FFRENLAFPQ G.EAREFPSE .......... ..QTRANSPA SR.......E
AC_RW_92RW   FFRENLAFQQ G.EARKFSPE Q......... ...TGANSPT SR.......E
AC_SE_SE94   FFRENLAFQQ G.EARKFSSE .......... ..QTGANSPT S.......RD
ACD_SE_SE8   FFREDMAFPQ G.KAREFPSE .......... ..QTRTNSPT S.......RE
```

```
ACG_BE_VI1   FFRENLAFQQ G.EARKFSSE .......... ..QTRANSPT SRANSPTSRE
AD_SE_SE69   FFRENLAFQQ G.KAREFPSE .......... ..QTRANSPS SR.......E
AD_SE_SE71   FSRENLAFQQ G.EARKFSSE .......... ..QTGTNSST SR......N.
ADHK_NO_97   FFRENLAFQQ R.KARELSSE .......... ..QTGAISPT SR.......E
ADK_CD_MAL   FFRENLAFPQ G.KAREFPSE .......... ..QTRANSPT SR.......E
AG_BE_VI11   FFRENLAFQQ G.EARKFSSE .......... ..QTGANSPT S.......RE
AG_NG_92NG   FFRENLAFQQ G.EAREFS.. .......... SEQARANSPT RR.......E
AGHU_GA_VI   FFRENLAFQQ G.EAREFS.. .......... PEQTRANSPT SR.......E
AGU_CD_Z32   FFRENLAFQQ G.EAREFSSE .......... ..QTRANSPT RR.......E
AJ_BW_BW21   FFRENLAFQQ G.KAREFSPE .......... ..QTRANSPT SR.......E
B_AU_VH_AF   FFREDLALPQ G.KARELSSE .......... ..QTRANSPT RR.......E
B_CN_RL42_   FFREDLAFPQ G.KARELSSE .......... ..QTRANSPT RG.......E
B_DE_D31_U   FFREDLAFLQ G.KAREFSSA .......... ..QTRANSPT RR.......E
B_DE_HAN_U   FFREDLAFPQ G.EARKFSSE .......... ..QTRANSPT RR.......E
B_FR_HXB2_   FFREDLAFLQ G.KAREFSSE .......... ..QTRANSPT RR.......E
B_GA_OYI__   FFREDLAFPQ G.KAREFSSE .......... ..QTRANSPT SR.......E
B_GB_CAM1_   FFRENLAFPQ G.EAREFSSE .......... ..QTRANSPT RR.......E
B_GB_GB8_A   FFREDLAFPQ G.KAREFSPE QTRANS.... ..PTRADSPT RR.......E
B_GB_MANC_   FFREDLALPQ G.KAREFSSE .......... ..QTRANSPT RG.......E
B_KR_WK_AF   FFREDLAFPQ G.KAREFPSE .......... ..QTRAISPA RR.......E
B_NL_3202A   FFREDLAFPQ G.KAREFSSE .......... ..QTRANSPT RR.......E
B_TW_TWCYS   FFRENLAFPQ G.KARKFSSE .......... ..QTRANSPT RG.......E
B_US_BC_L0   FFREDLAFPQ G.KAREFSSE .......... ..QTRANSST RR.......E
B_US_DH123   FFRENLAFPQ G.KAREFSSE .......... ..QTRAISPT RR.......E
B_US_JRCSF   FFREDLAFLQ G.KAREFPSE .......... ..QTRANSPT RR.......E
B_US_MNCG_   FFREDLAFLQ G.KAEFS.SE .......... ..QNRANSPT RR.......E
B_US_P896_   FFRENLAFPQ G.KAREFSSE .......... ..QTRANSPT RR.......E
B_US_RF_M1   FFRENLAFPQ G.KARELSSE .......... ..QTRANSPT RR.......E
B_US_SF2_K   FFREDLAFLQ G.KAREFSSE .......... ..QTRANSPT RR.......E
B_US_WEAU1   FFREDLVFPK G.KAREFSSE .......... ..QTRTNSPT RR.......E
B_US_WR27_   FFREXPAFPH X.KARXFPSE .......... ..QTRAISPT SR.......E
B_US_YU2_M   FFREDLAFPQ G.KARKFSSE .......... ..QTRANSPI RR.......E
BF1_BR_93B   FFRENLAFPQ G.KAREFPSE .......... ..QTRANSPT SR.......E
C_BR_92BR0   FFRENLAFPQ .GEARKSSSE .......... ..QNRANSPT RR.......E
C_BW_96BW0   FFRENLAFPQ G.EAREFPSE Q......... ...TRANSPT SR.......E
C_BW_96BW1   FFRENLAFPQ G.EAREFPSE .......... ..QTGANSPT SR.......K
C_BW_96BW1   FFRENLAFPQ G.EAREFPSE QTRAIS.... ........PT SR.......E
C_BW_96BW1   FFREDLAFPQ G.KAREFPSE QN........ ....RANSPT SR.......E
C_ET_ETH22   FFRETLAFQQ G.KAREFPSE QTRANSPTRE S.QTRANSPT TR.......E
C_IN_93IN1   FFRENLAFPQ G.EAREFPPE .......... ..QTGANSPT SR.......E
C_IN_93IN9   FFRENLAFPQ G.EAREFPPE .......... ..QTRADSPT SR.......E
C_IN_93IN9   FFRENLAFPQ G.EAREFPSE QTRANSPSS. ..QTRANSPS SR.......E
C_IN_94IN1   FFRENLAFPQ G.EAREFPPE .......... ..QTRANSPT SR.......E
C_IN_95IN2   FFRENLAFPQ G.EAREFPP. .......... ..ETRANSST SR.......E
CRF01_AE_C   FFRENLASQQ G.EAREFSSE .......... ..QTRANSPT SR......E.
CRF01_AE_C   FFRENLAFQQ G.EARKFPSE .......... ..QTRANSPT NG......E.
CRF01_AE_C   FFRENLAFQQ G.EAREFSSE .......... ..QTRANSPT SR......E.
CRF01_AE_T   FFREILAFQQ G.KAGKFSSE .......... ..QTRANSPA SR......K.
CRF01_AE_T   FFRENLAFQQ R.KAGEFSSE .......... ..QTRANSPT SR........
CRF01_AE_T   FFRENLAFQQ G.KAREFSSE .......... ..QTGANSSA SR......K.
CRF01_AE_T   FFRENLAFQQ G.KAGKFSSE .......... ..QTRANSPT SR......E.
CRF01_AE_T   FFRENLAFQQ G.KAGEFSSE .......... ..QTRANSPT SR......K.
CRF01_AE_T   FFRENLAFQQ G.KAGKFSSE .......... ..QTRTNSPT SR......K.
CRF02_AG_F   FFRENLAFQQ G.EARKFSSK .......... ..QTGTNSPT S.......RE
CRF02_AG_F   FFRENLAFQQ G.EARKFSSK .......... ..QTGTNSPT S.......RE
CRF02_AG_G   FFRENLAFQQ R.EARELSSE .......... ..QTGAISPT G.......RE
CRF02_AG_N   FFRENLAFQQ G.EARKFSSE .......... ..QTGTNSST S.......RE
CRF02_AG_S   FFRENLAFQQ G.EARKFSSE .......... ..QTGTNSPA S.......RE
CRF02_AG_S   FFRENLAFQQ G.EARKLSSE .......... ..QTGTNSPT S.......RE
CRF03_AB_R   FFRENLAFQQ R.EARKFSSE .......... ..QTRAISPT S.......RK
CRF03_AB_R   FFRENLAFQQ R.EARKFSSE .......... ..QTRAISPT S.......RK
```

```
CRF04_cpx_  FFRENVAFQQ R.EARKFSSE .......... ..QARANSPA RG.......M
CRF04_cpx_  FFRENVAFQQ R.KAGEFSSE .......... ..QARANSPT RR.......E
CRF04_cpx_  FFRENVAFQQ G.EARKFSSE .......... ..QDRANSPA RR.......E
CRF05_DF_B  FFRESLAFPQ G.EARELPPE .......... ..QTGALSPA SR.......E
CRF05_DF_B  FFRESLAFPQ G.KAREFPPE .......... ..QARTLSPT SR.......E
CRF06_cpx_  FFRENLAFQQ G.EAREFS.. .......... SEQARANSPT HR.......E
CRF06_cpx_  FFREDLAFQQ G.EARKFS.. .......... SEQARANSPT RG.......E
CRF06_cpx_  FFRENLAFQQ G.EAGELS.. .......... SEQARANSPT RR.......E
CRF06_cpx_  FFRENLAFPQ G.EAREFSPE QAR....... TEQARTLSPT CR.......E
CRF11_cpx_  FFRENLAFQQ R.KARELSPE .......... ..QTRANSPT SR.......E
CRF11_cpx_  FFRENLAFQQ G.EAREFPTE .......... ..QARANSPT SR.......E
D_CD_84ZR0  FFRENLAFPQ G.KAGELSSE .......... ..QTRANSPT S........R
D_CD_ELI_K  FFRENLAFPQ G.KAGELSPK .......... ..QTRANSPT SR.......E
D_CD_NDK_M  FFREDLAFPQ G.KAGEFSSE .......... ..QTRANSPT SR.......E
D_UG_94UG1  FFRENLAFPQ W.KAREFPSE QT........ ..PSRANSPT SR.......D
F1_BE_VI85  FFRENLAFQQ G.EARKFPSE .......... ..QTRANSPT SR.......E
F1_BR_93BR  FFRENLAFQQ G.EARKLHPE .......... ..QARAVSPA SR.......E
F1_FI_FIN9  FFRENLAFQQ G.EARKFPS. .......... ..ETRANSPA SR.......E
F1_FR_MP41  FFRENLAFQQ G.EARKFSSE .......... ..QARANSPA SG.......E
F2_CM_MP25  FFRENVAFQQ G.EARKFSSE .......... ..QTRANSPA SR.......E
F2KU_BE_VI  FFRENLAFQQ R.EAGKFSSE .......... ..QTRANSPT SR.......E
G_BE_DRCBL  FFRENLAFQQ G.EAREFP.. .......... SEQARANSPT RR.......E
G_NG_92NG0  FFRENLAFQQ G.EARKLS.. .......... PEQDRANSPT SR.......E
G_SE_SE616  FFRENLAFQQ G.EAREFS.. .......... SEQDRTNSPT CR.......K
H_BE_VI991  FFRENLAFQQ G.KAREFP.. .......... PEEARANSPT SR.......E
H_BE_VI997  FFRENLAFQQ R.EARKFS.. .......... PEQARANSPT SR.......E
H_CF_90CF0  FFRENLAFQQ R.EARKFS.. .......... PEQARTNSPT SR.......E
J_SE_SE702  FFREDLAFQQ R.EAREFSPE .......... ..QTRANSPT SR.......E
J_SE_SE788  FFREDLAFQQ R.EARELSPE .......... ..QTRANSPT SR.......E
K_CD_EQTB1  FFREVLASQQ R.EARKFSSE .......... ..QTRANSPT SR.......E
K_CM_MP535  FFRENLAFPQ G.EAREFSSE .......... ..QTRANSPT SR.......E
N_CM_YBF30  FFREELVSLQ R.ETRKLPPD NN........ ..KERAHSPA TR.......E
O_CM_ANT70  FFRQILASGG H.EARQLCAE T......... ...STPISPT DG.......G
O_CM_MVP51  FFREVLASGG H.EARQLCAE T......... ...SVPISPT NG.......G
O_SN_99SE_  FFREILASGG H.EARQLCAE T......... ...SVPISPT DD.......G
O_SN_99SE_  FFREILASGG H.EARQLCTE T......... ...SVPISPT DD.......G
U_CD___83C  FFRENLAFQQ G.EAREFSSE .......... ..QTRANSPT SR.......E

            51                                                  100
00BW0762_1  LQVR...... ......GDK. ....PHSEAG AERQ...... GTLNFPQITL
00BW0768_2  LQVRG..... ......DNN. ....PCSEAG AERQ...... GTLNCPQITL
00BW0874_2  PQARAISPTS REPQVRRDN. ....SRFEAG VEREG..... .TLNFPQITL
00BW1471_2  LQVR...... ......GDN. ....PRSEAG AERQG..... .TLNLPQITL
00BW1616_2  LQVR...... ......GNN. ....PLSEAG GERQ...... GTLNFPQITL
00BW1686_8  LQVR...... ......GDN. ....PRSEAG AERQ...... GTLNLPQITL
00BW1759_3  LQVRG..... .......NN. ....PRSEAG AERQ...... GNLNFPQITL
00BW1773_2  LQVR...... ......GDN. ....PRSEAG AERQ...... GTLNFPQITL
00BW1783_5  LQVR...... ......GDN. ....PCSEAG DERQ...... GTFNFPQITL
00BW1795_6  LQVR...... ......GDN. ....PLSEAG AERQ...... GTLNFPQITL
00BW1811_3  LQVR...... ......GDN. ....PRFEAG EKRQG..... .NLNFPQITL
00BW1859_5  LQVR...... ......GDD. ....PRSEAG AERQ...... GTLNFPQITL
00BW1880_2  LQVR...... ......GDN. ....PRSEAG AEGQ...... GTLNFPQITL
00BW1921_1  LQVR...... ......GDN. ....PCSEAG AERQG..... .TLNFPQITL
00BW2036_1  LQVR...... ......GDN. ....PRSEAG AERQ...... GTLNFPQITL
00BW2063_6  L..R...... ......GDN. ....PCSEAG DERQ...... GTLNFPQITP
00BW2087_2  NSPTSREL.. ...QVRGDN. ....PSIKAG PERQ...... GALNFPQITL
00BW2127_2  LQVR...... ......GDN. ....PRSEAG AERQG..... .SLNFPQITL
00BW2128_3  LQVR...... ......GDN. ....TRSEAG AKKQ...... GTLNFPQITL
00BW2276_7  LQVR...... ......GDN. ....PRAEAG AERQG..... .TLNFPQITL
00BW3819_3  LQVR...... ......GDN. ....PRSEAG DERQG..... .ALNFPQITL
00BW3842_8  LQVR...... ......GDN. ....PRSEAG AERQGT..LQ GTLNFPQITL
```

```
00BW3871_3  LQVR..... ......GDN. ....PRSEAG ADRQ..... GTLNFPQITL
00BW3876_9  LQVR..... ......GDN. ....PHSEAG AERQ..... GTLNFPQITL
00BW3886_8  LQVR..... ......GDN. ....PRSEAG AERQG..... .SLNFPQITL
00BW3891_6  LQVR..... ......GDN. ....PRSEAG AERQG..... .TLNFPQITL
00BW3970_2  LQVR..... ......GDN. ....PRSETG AEGQG..... .TFNFPQITL
00BW5031_1  LQVR..... ......GDN. ....PRSEAG DEREG..... .TLNFPQITL
 96BW01B21  LQVR..... ......GDN. ....PRSEAG AEGQG..... .ALNLPQITL
  96BW0407  LQVR..... ......GDN. ....PRSETR VEGQG..... .NFNFPQITL
  96BW0502  LQVR..... ......GDN. ....PRSEAG AEGQGT..LQ GTLNCPQITL
 96BW06_J4  LQIR..... ......GDN. ....PRFEAG TKRQ..... GTLNFPQITL
 96BW11_06  LRG..... .......NN. ....PCSEAG DERQ..... GTLNFPQITL
  96BW1210  LQVR..... ......GDN. ....PCSEAG AEGQG..... TTFSFPQITL
 96BW15B03  LQVR..... ......GDN. ....PRSEAG AERQ..... GTLNFPQITL
 96BW16_26  LQVW..... ......GDN. ....PRSETG AKGQ..... GTFNFPQITL
 96BW17A09  LQVR..... ......GDN. ....PRSEAG AERQG..... .TLNFLQITL
 96BWMO1_5  L..R..... ......GDN. ....PCSEAG DERQGT..LQ GALNFPQITL
 96BWMO3_2  LQAR..... .....TNSP. ....TSREAG VEGQG..... .TLNFPQITL
98BWMC12_2  P..... ...QARGDN. ....TRFEAG DEGQG..... .TLNFPQITL
98BWMC13_4  P..R..... ......GDN. ....PCSEAG AERQ..... GTLNLPQITL
98BWMC14_a  LQVR..... ......GDN. ....PRSEAG AEGQ..... GTLNFPQITL
98BWMO14_1  LQVREQTR.. ...ANSSTS. ....RELQAG AKRQ..... GALNCPQITL
98BWMO18_d  LQVR..... ......GDN. ....PCSEAG AERQGS.... .TLNFPQITL
98BWMO36_a  LQVR..... ......GDK. ....PRSEAG AEGQG..... .TLNFPQITL
98BWMO37_d  LQVR..... ......GDN. ....PRSEAG GERQG..... .TLKFPQITL
99BW3932_1  FQVR..... ......GDN. ....PCSEAG AERQG..... .SLNFPQITL
99BW4642_4  LQIR..... ......GDD. ....PRSEAG AERQ..... RTLNFPQITL
99BW4745_8  LQVR..... ......GGN. ....PHSEAG AERQG..... .TLNFPQITL
99BW4754_7  LQVR..... ......GDN. ....PHSEAR VKGQ..... GTPNFPQITL
99BWMC16_8  LQVR..... ......GDK. ....SRSEAG VEKQG..... .NLNFPQITL
A2_CD_97CD  ...GGR.... .......DN. ....LLAEAG E..QG...AV HPCNFPQITL
A2_CY_94CY  LENGGR.... .......DN. ....LLPEAG TGDQG...TI QSCNFPQITL
A2D___97KR  LWNGGG.... .......DN. ....PLAEAG AEKQG...TT HSCNFPQITL
A2G_CD_97C  PRVRR..... ......GDS. ....LLPEAG DEG...KGAV YPCNFPQITL
A_BY_97BL0  LD.GGR.... .......DN. ....PLPETG TERQG...TV SSFNFPQITL
A_KE_Q23_A  LWDGGR.... .......DS. ....LPSEAG AERQGT..G. PTLSFPQITL
A_SE_SE659  PWDRRR.... .......DS. ....LPSETG ADP..... .TFSFPQITL
A_SE_SE725  FWDGGR.... .......DS. ....LPSEAG AERQGT..E. LTFSFPQITL
A_SE_SE753  LWNEGR.... .......DS. ....LPSEAG AEG..T..R. PTFSFPQITL
A_SE_SE853  LWDGGS.... .......DN. ....LPSEAG AERQGT..G. PTLSFPQITL
A_SE_SE889  LWDGGR.... .......DN. ....LPSEAG EERQGV..GG TTLNFPQITF
A_SE_UGSE8  ..DGGR.... .......DS. ....LPSEAG AKQP..... .TFSFPQITL
A_UG_92UG0  LWDEGR.... .......DS. ....LPSEAG AERQGP..E. PTFSFPQITL
A_UG_U455_  LWDGGK.... .......DD. ....LPCETG AERQ....GT DSFSFPQITL
AC_IN_2130  LQIR..... ......GDN. ....PRTEAG AKRQG..... .TLNFPQITL
AC_RW_92RW  LWNGG..... ......RDS. ....LSSETG AERQG..... .TFNFPQITL
AC_SE_SE94  LRDGGR.... .......D.. .....NSEAG TDRQGT..G. PAFSFPQITL
ACD_SE_SE8  LRVWRR.... .......DN. ....PLPEAG AERQ....GT VSFSLPQITL
ACG_BE_VI1  LWEGGR.... .......DR. ....LLPEAG TEGQG...TI SSFNFPQITL
AD_SE_SE69  LRVWRG.... ......DS.. ....TFSETG AER....QGA VSFSFPQITL
AD_SE_SE71  LWDGGR.... .......DS. ....LPSEAG AEKQG...TG STLNFPQITL
ADHK_NO_97  LWDRGR.... ......DN.. ....LLSEAG TEGQG...TA PSLSFPQITL
ADK_CD_MAL  LRVWGG.... ......DK.. ....TLSETG AERQ....GI VSFSFPQITL
AG_BE_VI11  LGDGGR.... .......DN. ....PLSEAG TEGHG...TI SSLNFPQITL
AG_NG_92NG  LRVRR..... ......GDS. ....PFPEAG AEG...KGIT SIN.LPQITL
AGHU_GA_VI  LRVRR..... ......GDS. ....PLPEAG AKG....KGA VSFNLPQITL
AGU_CD_Z32  LRDER..... ......GDN. ....LLSEAG TEGQG...TI PSFSFPQITL
AJ_BW_BW21  LRVWR..... ......GDS. ....PLPEAG GEGQ....GT VSFNFPQITL
B_AU_VH_AF  LQVWGR.... ......DNN. ....SLSEAG ADR....QGT VSFSFPQITL
B_CN_RL42_  LQVWGR.... ......DNN. ....SISEAG ADR....QGT ISFSFPQITL
B_DE_D31_U  LQVWGR.... ......DSN. ....SLSEAG ADR....QGT VSFSFPQITL
B_DE_HAN_U  LQVWG..... ......SNS. ....SLSEAG ADR....QGT VSLSLPQITL
```

```
B_FR_HXB2_  LQVWGR.... ......DNN. ....SPSEAG ADR....QGT VSFNFPQVTL
B_GA_OYI__  LRVWGR.... ......DNN. ....SPSEAG ADR....QGT VSFNLPQITL
B_GB_CAM1_  LQVWGR.... ......ENN. ....SLSEAG ADR....QGT VSFSFPQITL
B_GB_GB8_A  LQVRGR.... ......DNN. ....SLTETG ADK....QGT VSFSFPQITL
B_GB_MANC_  LQVWGR.... ......DNN. ....SCSEAG TDR....QGT VSLSFPQTTL
B_KR_WK_AF  LQVWGR.... ......DNN. ....SLSEAG ANR....QGT VSFSFPQITL
B_NL_3202A  LQVWGR.... ......DNN. ....SLSEAG AEG....QGT VSLSLPQITL
B_TW_TWCYS  LQVWGT.... ......DNN. ....SLSEAG ADR....QGP VSFSFPRITL
B_US_BC_L0  LQVWGG.... ......DNN. ....SPSEAG AGR....QGN VSLSFPQITL
B_US_DH123  LQVWGG.... ......DSN. ....SLSEAG AEG......T ISLSLPQITL
B_US_JRCSF  LQVWGR.... ......DSN. ....SLSEAG AEAGADRQGI VSFNFPQITL
B_US_MNCG_  LQVWGR.... ......DNN. ....SLSEAG EEAGDDRQGP VSFSFPQITL
B_US_P896_  LQVWGG.... ......DNN. ....SLSEAG ADR....QGT VSLSFPQITL
B_US_RF_M1  LQVWGR.... ......DN.. ....SLSEAG EDR....QGT VSFSFPQITL
B_US_SF2_K  LQVWGG.... ......ENN. ....SLSEAG ADR....QGT VSFNFPQITL
B_US_WEAU1  LQVQGR.... ......DNN. ....SLSEAG ANR....QGA VSFNFPQITL
B_US_WR27_  LQVWXR.... ......DNN. ....SLSEAG TDR....QGT VSFSFPQITL
B_US_YU2_M  RQVWRR.... ......DNN. ....SLSEAG ADR....QGT VSFSFPQITL
BF1_BR_93B  LQVWGR.... ......GNN. ....SLSETG ADR....QGD VSFGFPQITL
C_BR_92BR0  LQVWGR.... ......DNN. ....SLSEAG DDRQG..... TALNFPQITL
C_BW_96BW0  LQVR...... ......GDN. ....PRSETR AEGQG..... .TLNFPQITL
C_BW_96BW1  LRG....... .......NN. ....PCSEAG DERQ...... GTLNFPQITL
C_BW_96BW1  LQVR...... ......GDN. ....PCSEAG AEGQG..... TTFSFPQITL
C_BW_96BW1  LQVR...... ......GDN. ....PRSEAG AERQ...... GTLNFPQITL
C_ET_ETH22  LQVR...... ......GSN. ....TFSEAG AERQG..... .SLNFPQITL
C_IN_93IN1  LQVR...... ......GDN. ....PSSKAG AERQG..... .TLNFPQITL
C_IN_93IN9  LQVR...... ......GDT. ....PSSKAG AERQG..... .TLNFPQITL
C_IN_93IN9  LQVR...... ......GDN. ....PRSEAG AKRQG..... .TLNFPQITL
C_IN_94IN1  LQVR...... ......GDT. ....PSSKAG AEREG..... .TLNFPQITL
C_IN_95IN2  LQVR...... ......GDN. ....PSSEAG AERQG..... .TFNFPQITL
CRF01_AE_C  LGDGGR.... .......DN. ....LLPEAG AERQG...TP FSFSFPQITF
CRF01_AE_C  LGDGGR.... .......DN. ....LLPEAG AERQE...TA SSFSFPQITL
CRF01_AE_C  LRDGGR.... .......DN. ....LLLEAG AERQG...TS SSLSFPQITL
CRF01_AE_T  LGDGGR.... .......DN. ....LLTEAG AERQG...TS SSFSFPQITL
CRF01_AE_T  MGDGGR.... .......DN. ....LLTEAG AERQ....GS SSFSFPQITL
CRF01_AE_T  LGDGG..... .......... .........G AERQG...TS SSFSFPQITL
CRF01_AE_T  LGDGGR.... .......DN. ....LLAETG AERQG...TP SSFNFPQITL
CRF01_AE_T  LGDGGR.... .......DNG GRDNLLTEAG AERQG...TS SSFSFPQITL
CRF02_AG_F  LGDGGR.... .......DN. ....LLPEAG AERQG...TP SSFSFPQITL
CRF02_AG_F  LWDGGR.... .......DN. ....LPSEAG SEGPG...TI SSLSFPQITL
CRF02_AG_G  LWDKGR.... .......DN. ....LPSEAG TEGPG...TI SPSSFPQITL
CRF02_AG_N  LWDGGR.... .......NN. ....LLSAAG TEGQG...TI SSFNFPQITL
CRF02_AG_S  PWDRGR.... .......DT. ....SLSTAG TEGQG...AI SSFNFPQITL
CRF02_AG_S  LWDGGR.... .......DN. ....LLSEAG TGGQG...TI SSLSFPQITL
CRF03_AB_R  LWDGGR.... .......DN. ....LLPEAG TGGQG...TI PSFNFPQITL
CRF03_AB_R  LWDGGR.... .......DN. ....PLPETG TEGQG...TA SSFNFPQITL
CRF04_cpx_  LREERG.... ......DN.. ....PLPETG TERQG...TA SSFNLPQITL
CRF04_cpx_  LRDERG.... ......DN.. ....LLSEAG TEGQ....GT ISFNFPQITL
CRF04_cpx_  LRDERG.... ......DN.. ....LLSEAG TEGQ....GT ISFNFPQITL
CRF05_DF_B  LQVWGG.... ......DS.. ....LLSEAG TEGQ....GT ISHNFPQITL
CRF05_DF_B  LRVWRG.... ......DN.. ....PLAEAG AEG...RGTV PSLSFPQITL
CRF06_cpx_  LRFRR..... ......GDS. ....PLPETG VEGEGGKGAI SLS.LPQITL
CRF06_cpx_  LRVRR..... ......GDS. ....PLPGAE AEG...KGAI SLN.FPQITL
CRF06_cpx_  LRVRR..... ......GDS. ....PLPEAG TEGKG.KGAI SLS.FPQITL
CRF06_cpx_  LRVRG..... ......GNS. ....PLPEAG AEG...EGAI SLS.FPQITL
CRF11_cpx_  LRVRR..... ......GDS. ....PLPETG AEGEGE..GA ISFNLPQITL
CRF11_cpx_  LRVRG..... ......GDN. ....PLPETG AQGE....GT ISYNFPQITL
D_CD_84ZR0  LRVWGG.... ......DN.. ....PLPETG TEG..QRQGT VSLSFPQITL
D_CD_ELI_K  LRVWGR.... ......DN.. ....PLSKTG AE....RQGT VSFNFPQITL
D_CD_NDK_M  LRVWGG.... ......DN.. ....PLSETG AE....RQGT VSFSFPQITL
```

```
D_UG_94UG1  LRIRGG.... ......DN.. ....TSSETG AER....QGT VSFNLPQITL
F1_BE_VI85  LRVQRG.... ......DN.. ....PLSEAG AERR...GTV PSLSFPQITL
F1_BR_93BR  LQVRGG.... ......DN.. ....PISEAG AERR...GTV PSLSFPQITL
F1_FI_FIN9  PRDQRR.... ......GN.. ....SLSEAG AERR...GTV PSLSFPQITL
F1_FR_MP41  LRVQRG.... ......NN.. ....PLSEAG AEGRGT.GTV SSLSLPQITF
F2_CM_MP25  LRVRGG.... ......DS.. ....SLPEAG AERQG...TG SSLDFPQITL
F2KU_BE_VI  LRVWGG.... ......DK.. ....PLSEAG DERQG...TG ASFNLPQITL
G_BE_DRCBL  LRVRG..... ......GDS. ....PLPEAG AEG...KGTI S.SIFPQITL
G_NG_92NG0  LRIRR..... ......GDS. ....PLPEAG AKG...EGAI SLN.FPQITL
G_SE_SE616  PRVRR..... ......GDS. ....PLPEAG DEG...KGAI S...LPQITL
H_BE_VI991  LRVRR..... ......GDH. ....PLSEAG AE.....RTG TSFNFPQITL
H_BE_VI997  LRVRG..... ......GDD. ....LLPEAG AE.....GQG TSLCFPQITL
H_CF_90CF0  LRVRR..... ......GDD. ....PLSEAG AAE....GQG TSLSFPQITL
J_SE_SE702  PRVRR..... ......GD.. ....PLPETG AEGQ....GT VSSNFPQITL
J_SE_SE788  PRARR..... ......GD.. ....PLPETG AEGQ....GT VSSNFPQITL
K_CD_EQTB1  LWVRGE.... ......DN.. ....PLSETG NERSG...TG SSFNFPQITL
K_CM_MP535  LRVRGG.... ......DN.. ....PLSEAG DQRQG...TE PSFNFPQITL
N_CM_YBF30  LWVSG..... ......GEEH TGEGDAGEPG EDRE...LSV PTFNFPQITL
O_CM_ANT70  GSEGTG.... ......ESG. ......TERG PER......A LSVCLPQIPL
O_CM_MVP51  GSEGTR.... ......ESE. ......SEGG SGR......A VPICLPQIPL
O_SN_99SE_  GNEGTR.... ......ESG. ......TEGG PER......T LSVCLPQIPL
O_SN_99SE_  GNEGTR.... ......KSG. ......TEGG PER......T LSVCLPQIPL
U_CD___83C  LRVRR..... ......GDN. ....PLAEAG AEGQG..VPL PSFNFPQITL

            101                                                   150
00BW0762_1  WQRPLVSIKV GGQIKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
00BW0768_2  WQRPLVSIRV GGQIKEALLD TGADDTVLEE ISLPGKWKPK MIGGIGGFIK
00BW0874_2  WQRPLVSIKI EGQIREALLD TGADDTVLEE ITLSGRWKPK MIGGIGGFIK
00BW1471_2  WQRPLVSIKV GGQIKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
00BW1616_2  WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
00BW1686_8  WQRPLVTIKV GGQVKEALLD TGADDTVLEE ISLPGKWKPK MIGGIGGFIK
00BW1759_3  WQRPLVSIKV GGQVKEALLD TGADDTVLEE LALPGRWKPK MIGGIGGFIK
00BW1773_2  WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
00BW1783_5  WQRPLVTIKV GGQIKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
00BW1795_6  WQRPLVSIKV GGQIKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
00BW1811_3  WQRPLVTIKV GGQTKEALLD TGADDTVLEE MNLPGKWKPK MIGGIGGFIK
00BW1859_5  WQRPLVSIKV GGQIKEALLD TGADDTVLEE ISLPGKWKPK MIGGIGGFIK
00BW1880_2  WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
00BW1921_1  WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
00BW2036_1  WQRPLVSIKV GGQIKEALLD TGADDTVLEE MSLPGKWKPK MIGGIGGFIK
00BW2063_6  WQRPLVSIKV GGQIKEALLD TGADDTVLED INLLGKWKPK MIGGIGGFIK
00BW2087_2  WQRPLVSIKI EGQIKEALLD TGADDTVLEE MNLPGRWKPK MIGGIGGFIK
00BW2127_2  WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
00BW2128_3  WQRPLVSIKI GDQVKEALLD TGADDTVLEE IKLPGKWKPK MIGGIGGFIK
00BW2276_7  WQRPLVSIKV GGQIREALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
00BW3819_3  WQRPLVSIKV GGQIKEALLD TGADDTVLED INLSGKWKPK MIGGIGGFIK
00BW3842_8  WQRPLVTIKV GGQIKEALLD TGADDTVLEE MNMPGKWKPK MIGGIGGFIK
00BW3871_3  WQRPLVTIKV GGQVKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
00BW3876_9  WQRPLVSIKV GGQIKEALLD TGADDTVLEE ISLPGKWKPK MIGGIGGFIK
00BW3886_8  WQRPLVSIKV GGQIKEALLD IGADDTVLEE LSLPGKWKPK MIGGIGGFIK
00BW3891_6  WQRPLVTVKV GGQIKEALLD TGADDTVLEE ISLPGKWKPK MIGGIGGFIK
00BW3970_2  WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
00BW5031_1  WQRPLVSIKV GGQIKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
 96BW01B21  WQRPLVTIRV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
  96BW0407  WQRPLVSIKV GGQIREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
  96BW0502  WQRPLVSIKV GGQIKEALLD TGADNTVLEE INLPGKWKPK MIGGIGGFIK
 96BW06_J4  WQRPLVSIKV GGQIKEALLD TGADDTVLEE MSLSGKWKPK MIGGIGGFIK
 96BW11_06  WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWRPK MIGGIGGFIK
  96BW1210  WQRPLVSIKV GGQIKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
 96BW15B03  WQRPLVSIKV GGQIKEALLD TGADDTVLEE MSLPGKWKPK MMGGIGGFIK
 96BW16_26  WQRPLVSIKV GGQVKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
```

```
96BW17A09  WQRPLVSIKV GGQIREALLD TGADDAVLED INLPGKWKPK MIGGIGGFIK
96BWMO1_5  WQRPLVSIKV GGQIKEALLD TGADDTVLEE VNLPGKWKPK MIGGIGGFIK
96BWMO3_2  WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
98BWMC12_2 WQRPLVSIKI GGQIREALLD TGADDTVLEE LSLPGRWKPK MIGGIGGFIK
98BWMC13_4 WQRPLVSIKV GGQVKEALLD TGADDTVLED IELPGKWRPK MIGGIGGFIK
98BWMC14_a WQRPLVSIKV GGQTKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
98BWMO14_1 WQRPLVSIKI GGQIKEALLD TGADDTVLEE MSLPGRWKPK MIGGIGGFIK
98BWMO18_d WQRPLVSIKV GGQIKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
98BWMO36_a WQRPLVSIKV GGQTKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
98BWMO37_d WQRPLVSIKV GGQIREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
99BW3932_1 WQRPLVPIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
99BW4642_4 WQRPLVSIKV GGQIKEALLD TGADDTVLED ISLPGKWKPK MIGGIGGLIK
99BW4745_8 WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
99BW4754_7 WQRPLVSIKV GGQIKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
99BWMC16_8 WRRPLVTIKV GGQIKEALLH PGADDTVLEE INLPRKWKPK MIGGIGGFIK
A2_CD_97CD WQRPLVTVKI EGQLREALLD TGADDTVLED INLPWKWKPK MIGGIGGFIK
A2_CY_94CY WQRPLVTVKI EGQLKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
A2D___97KR WQRPLVTVKI EGQLREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
A2G_CD_97C WQRPLVTVKI GGQLIEALLD TGADDTVLED INLPGRWKPK MIGGIGGFIK
A_BY_97BL0 WQRPLVTVRI GGQLKEALLD TGADXTVLED INLPGKWKPK MIXGIXGFIK
A_KE_Q23_A WQRPLVTVRI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
A_SE_SE659 WQRPLVTVKV GGQLREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
A_SE_SE725 WQRPLVTVKI GGQLREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
A_SE_SE753 WQRPLVTVKI EGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
A_SE_SE853 WQRPLVTVKV GGQLKEALLD TGADDTVLED INLPGKWRPK MIGGIGGFIK
A_SE_SE889 WQRPLVTVRI GGMQKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
A_SE_UGSE8 WQRPIVTVRI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
A_UG_92UG0 WQRPLVTVKI GGQLKKALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
A_UG_U455_ WQRPLVTVKI GGQLIEALLD TGADDTVLED INLPGKWKPK IIGGIGGFIK
AC_IN_2130 WQRPLVSIRV GGQTKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
AC_RW_92RW WQRPLVTVKI GGQLREALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
AC_SE_SE94 WQRPLVTVRI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
ACD_SE_SE8 WQRPLVKVKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
ACG_BE_VI1 WQRPLVTVRL GGQLIEALLD TGADDTVLEQ INLPGKWKPK MIGGIGGFIK
AD_SE_SE69 WQRPLVTVKI GGQLREALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
AD_SE_SE71 WQRPLVTVKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
ADHK_NO_97 WQRPVVTVKV GGQLKEALLD TGADDTVLED MNLPGKWKPK MIGGIGGFIK
ADK_CD_MAL WQRPVVTVRV GGQLKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
AG_BE_VI11 WQRPLVTVRI GGQLIEALLD TGADDTVLAE ISLPGKWKPK MIGGIGGFIK
AG_NG_92NG WQRPLVTVRI GGQLIEALLD TGADDTVLEQ INLPGKWKPK MIGGIGGFIK
AGHU_GA_VI WQRPLVTVKI GGQLIEALLD TGADDTVLEE INLLGKWKPK MIGGIGGFIK
AGU_CD_Z32 WQRPIVTVKI GGQPIEALLD TGADDTVLEE IKLPGKWKPK MIGGIGGFIK
AJ_BW_BW21 WQRPLVTIRV AGQVKEALLD TGADDTVLEE MELPGKWKPK MIGGIGGFIK
B_AU_VH_AF WQRPIVTIKI GGQLKEALLD TGADDTVLEE MCLPGRWKPK MIGGIGGFIK
B_CN_RL42_ WQRPLVTIKV GGQLKEALLD TGADDTVLED MNLPGRWKPK MIGGIGGFIK
B_DE_D31_U WQRPLVTIKI GGQLKEALLD TGADDTVLEE MSLPGRWKPK MIGGIGGFIK
B_DE_HAN_U WQRPLVTIKI GGQLKEALLD TGADDTVVEE MSLPGRWKPK MIGGIGGFIK
B_FR_HXB2_ WQRPLVTIKI GGQLKEALLD TGADDTVLEE MSLPGRWKPK MIGGIGGFIK
B_GA_OYI___ WQRPIVTIKI GGQLKEALLD TGADDTVLEE MNLPGRWKPK MIGGIGGFIK
B_GB_CAM1_ WQRPLVTIKI GGQLKEALLD TGADDTVLEE MNLPGRWKPK MIGGIGGFIK
B_GB_GB8_A WQRPIVTIKI GGQLKEALLD TGADDTVLED MNLPGRWKPK MIGGIGGFIK
B_GB_MANC_ WQRPLVTIKI GGQLKEALLD TGADDTVLEE MNLPGRWKPK MIGGIGGFIK
B_KR_WK_AF WQRPLVAIKI GGQLKEALLD TGADDTVLEE MSLPGRWKPK MIGGIGGFIK
B_NL_3202A WQRPLVTIKI GGQLKEALLD TGADDTVLEE MNLPGRWKPK MIGGIGGFIK
B_TW_TWCYS WQRPLVTIRI GGQLKEALLD TGADDTVLEE MNLPGKWKPK MIGGIGGFIK
B_US_BC_L0 WQRPLVTIKI GGQLKEALLD TGADDTVLEE MNLPGRWKPK MIGGIGGFIK
B_US_DH123 WQRPLVKIKI GGQLKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
B_US_JRCSF WQRPLVTIKI GGQLKEALLD TGADDTVLED MDLPGRWKPK MIGGIGGFIK
B_US_MNCG_ WQRPIVTIKI GGQLKEALLD TGADDTVLGE MNLPRRWKPK MIGGIGGFIK
B_US_P896_ WQRPLVTIKV GGQLKEALLD TGADDTVLED MSLPGRWKPK MIGGIGGFIK
B_US_RF_M1 WQRPIVTVKI GGQLKEALLD TGADDTVLEE MNLPGKWKPK MIGGIGGFIK
```

```
B_US_SF2_K  WQRPLVTIRI GGQLKEALLD TGADDTVLEE MNLPGKWKPK MIGGIGGFIK
B_US_WEAU1  WQRPLVTIKI EGQLKEALLD TGADDTVLED MNLPGKWKPK MIGGIGGFIK
B_US_WR27_  CQRPLVAIKI GGQIKEALLD TGADDTVLEE MSLPGRWKPK MVGGIGGFIK
B_US_YU2_M  WQRPLVTIKI GGQLKEALLD TGADDTVLEE MNLPGRWKPK MIGGIGGFIK
BF1_BR_93B  WQRPLVTVKI GGQLKEALLD TGADDTVLEE INLPGRWKPK MIGGIGGFIK
C_BR_92BR0  WQRPLVNIKV GGQLKEALLD TGADDTVLEE IKLPGNWKPK MIGGIGGFIK
C_BW_96BW0  WQRPLVSIKV GGQIREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
C_BW_96BW1  WQRPLVSIKV GGQIKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
C_BW_96BW1  WQRPLVSIKV GGQIKEALLD TGADDTVLED INLPGKWRPK MIGGIGGFIK
C_BW_96BW1  WQRPLVSIKV GGQIKEALLD TGADDTVLEE MSLPGKWKPK MMGGIGGFIK
C_ET_ETH22  WQRPLVTIKI GGQLKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
C_IN_93IN1  WQRPLVSIRV GGQIKEALLD TGADDTVLEE VNLPGKWKPK MIGGIGGFIK
C_IN_93IN9  WQRPLVSIRV GGQIKETLLD TGADDTVLEE VNLPGKWKPK MIGGIGGFIK
C_IN_93IN9  WQRPLVSIKV GGQIREALLD TGADDTVLEE VNLPGKWKPK MIGGIGGFIK
C_IN_94IN1  WQRPLVSIRV GGQTREALLD TGADDTVLEE VNLPGKWKPK MIGGIGGFIK
C_IN_95IN2  WQRPLVSIRV GGQIKEALLD TGADDTVLEE VNLPGKWKPK MIGGIGGFIK
CRF01_AE_C  WQRPLVTVKI EGQLKEALLD TGADDTVLED VSLPGKWRPK MIGGIGGFIK
CRF01_AE_C  WQRPLVTVKV GGQLKEALLD TGADDTVLED INLPGRWKPK MIGGIGGFIK
CRF01_AE_C  WQRPIVTVKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF01_AE_T  WQRPLVTVKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF01_AE_T  WQRPLVTVKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF01_AE_T  WQRPLVTIKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF01_AE_T  WQRPLVTVKI GGELKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF01_AE_T  WQRPLVTVKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF01_AE_T  WQRPLVTVKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF02_AG_F  WQRPLVTVKI GGQLIEALLD TGADDTVLED INLPGKWKPR MIGGIGGFIK
CRF02_AG_F  WQRPLVTVRI GGQLIEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
CRF02_AG_G  WQRPLVTVRI GGQLIEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
CRF02_AG_N  WQRPLVTVRI EGQLIEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF02_AG_S  WQRPLVTVKI GGQLIEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF02_AG_S  WQRPLFTVRI EGQLIKALLD TGANDTVLEK INLPGKWKPK MIGGIGGFIK
CRF03_AB_R  WRRPLVTVRI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF03_AB_R  WQRPLVTVRI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF04_cpx_  WQRPLVTIKL GGQIREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF04_cpx_  WQPPLVTIKI GGQIREALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
CRF04_cpx_  WQRPLVTIKI GGQLREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF05_DF_B  WQRPVVTIRI GGQLKEALLD TGADDTVLEE INLPGRWKPK MIGGIGGFIK
CRF05_DF_B  WQRPVVTIKI EGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF06_cpx_  WQRPLVTVRI GGQLIEALLD TGADDTVLEE MNLPGKWKPK MIGGIGGFIK
CRF06_cpx_  WQRPLVTVRI GGQLIEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF06_cpx_  WQRPLVTVRI GGQLIEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF06_cpx_  WQRPLVTVGI EGQLIEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
CRF06_cpx_  WQRPLVTVKV GEQLIEALLD TGADDTVLEE INLPGRWKPK MIGGIGGFIK
CRF11_cpx_  WQRPIVKIKV AGQLKEALLD TGADDTVLEE INLPGRWKPK MIGGIGGFIK
CRF11_cpx_  WQRPVVPVKV AGQLKEALLD TGADDTVLEE MSLPGRWKPK MIGGIGGFIK
D_CD_84ZR0  WQRPVVTIKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
D_CD_ELI_K  WQRPLVAIKI GGQLKEALLD TGADDTVLEE MNLPGKWKPK MIGGIGGFIK
D_CD_NDK_M  WQRPLVTIKI GGQLKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
D_UG_94UG1  WQRPVVTVKI GGQLKEALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
F1_BE_VI85  WQRPLVTIKI GGQIKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
F1_BR_93BR  WQRPLVTIRV GGQLKEALLD TGADDTVLED VNLPGKWKPK MIGGIGGFIK
F1_FI_FIN9  WQRPLVTIKI GGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
F1_FR_MP41  WQRPLVTIRV GGQLREALLD TGADDTVLED IDLPGKWKPK IIGGIGGFIK
F2_CM_MP25  WQRPVVTIKV GGQLREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
F2KU_BE_VI  WQRPIVTIKI GGQLREALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
G_BE_DRCBL  WQRPIVKVRI GGQLIEALLD TGADDTVLEE IDLPGKWKPK MIGGIGGFIK
G_NG_92NG0  WQRPLVTVKI GGQLIEALLD TGADDTVLEG INLPGKWKPK MIGGIGGFIK
G_SE_SE616  WQRPLVTVKI GGQLIEALLD TGADDTVLEE INLPGRWKPK MIGGIGGFIK
H_BE_VI991  WQRPIVTVKI EGQLKEALLD TGADDTVLED INLPGKWKPK MIGGIGGFIK
H_BE_VI997  WQRPLVTVKI EGQLREALLD TGADDTVLEE INLLGRWKPK MIGGIGGFIK
H_CF_90CF0  WQRPLVTVKI EGQLREALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
J_SE_SE702  WQRPLVTIRI GGQLREALLD TGADDTVLEE IDLPGKWKPK MIGGIGGFIK
```

```
J_SE_SE788   WQRPLVTIRI GGQLREALLD TGADDTVLED IDLPRKWKPK MIGGIGGFIK
K_CD_EQTB1   WQRPVVTVKV GGQLREALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
K_CM_MP535   WQRPIVTIKV GGQLREALLD TGADDTVLEE INLPGKWKPK MIGGIGGFIK
N_CM_YBF30   WQRPVITVKI GKEVREALLD TGADDTVIEE LQLEGKWKPK MIGGIGGFIK
O_CM_ANT70   WDRPIVTARV GGHLCEVLLD TGADDTVLNN IQLEGKWKPK MIGGIGGFIK
O_CM_MVP51   WDRPIVTAKV GGHLCEALLD TGADDTVLNN IQLEGRWTPK MIGGIGGFIK
O_SN_99SE_   WDRPVVTARV GGHLCEVLLD TGADDTVLTN IQLEGKWTPK MIGGIGGFIK
O_SN_99SE_   WDRPIVPARV GGHLCEVLLD TGADDTVLNN IQLEGKWTPK MIGGIGGFIK
U_CD___83C   WQRPLVTVKI GGQLKEALLD TGADDTVVEE MTLPGKWKPK MIGGIGGFIK

             151                                                 200
00BW0762_1   VRQYDQILIE ICGKRAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW0768_2   VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW0874_2   VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW1471_2   VRQYDQIVIE ICGKKAIGSV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW1616_2   VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW1686_8   VRQYDQISIE ICGKRAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW1759_3   VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW1773_2   VRQYDQISIE ICGKKAIGAV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW1783_5   VRQYDQILIE ICGKKAIGSV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW1795_6   VRQYDHIPIE ICGKRAIGTV LVGPTPVNII GRNMLTQLGC ILNFPISPIE
00BW1811_3   VRQYDEILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW1859_5   VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW1880_2   VRQYDQILIE ICGKRAMGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIK
00BW1921_1   VRQYDQITIE ICGKKAIGAV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW2036_1   VRQYDQIPIE ICGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW2063_6   VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW2087_2   VRQYDQILIE ICGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW2127_2   VRQYDQVVIE ICGKKTIGTV LVGPTPVNIV GRNLLTQLGC TLNFPISPIE
00BW2128_3   VRQYDEIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW2276_7   VRQYDQILIE ICGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW3819_3   VRQYEQVPIE ICGKKAIGTV LVGPTPANII GRNLLTQLGC TLNFPISPIE
00BW3842_8   VRQYDQIVIE ICGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW3871_3   VRQYEQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW3876_9   VRQYDQILVE ICGKKAIGTV SVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW3886_8   VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW3891_6   VRQYDQIAIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW3970_2   VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
00BW5031_1   VRQYDQIMIE ICGQKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
 96BW01B21   VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
  96BW0407   VRQYEQILIE ICGKKTIGTV LVGPTPVDII GRNMLTQLGC TLNFPISPIE
  96BW0502   VRQYDQIVIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
 96BW06_J4   VRQYDQIPIE ICGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
 96BW11_06   VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIE
  96BW1210   VRQYDQIVIE ICGKKAIGSV LVGPTPVNII GRNMLTQLGC TLNFPISPIK
 96BW15B03   VRQYDQILIE ICGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
 96BW16_26   VRQYDQITIE ICGKKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
 96BW17A09   VRQYDQIVIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
 96BWMO1_5   VRQYDQIPIE ICGKRAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
 96BWMO3_2   VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
98BWMC12_2   VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
98BWMC13_4   VRQYDRIPIE ICGKRAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
98BWMC14_a   VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
98BWMO14_1   VGQYDQIPIE ICGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
98BWMO18_d   VKQYEQILIE ICGKKAIGTV LVGPTPVNII GRNMLAQLGC TLNFPISPIE
98BWMO36_a   VRQYDQILIE ICGKKAIGTV LVGPTPINII GRNMLTQLGC TLNFPISPIE
98BWMO37_d   VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
99BW3932_1   VRQYDQVVIE ICEKKTIGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIE
99BW4642_4   VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
99BW4745_8   VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
99BW4754_7   VRQYDQIHIE ICGKRAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
```

```
99BWMC16_8  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTPLGC TLNFPISPIE
A2_CD_97CD  VRQYDQIVIE ICGKRAIGTV LVGPTPVNII GRNMLVQLGC TLNFPISPIE
A2_CY_94CY  VRQYDQIAIE ICGKRAIGTV LVGPTPVNII GRNMLVQLGC TLNFPISPIE
A2D___97KR  VRQYDQITIE ICEKRAIGTV LVGPTPVNII GRNMLVQLGC TLNFPISPIE
A2G_CD_97C  VRQYDQILIE ISGKKAIGTV LVGPTPINII GRNMLIQIGC TLNFPISPIE
A_BY_97BL0  VRQYDQILVE ICXKKAIXTV LVGPTPXNII XRNMLTQLGC TLNFPISPIE
A_KE_Q23_A  VKQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
A_SE_SE659  VKQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
A_SE_SE725  VKQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
A_SE_SE753  VKQYDQVLIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
A_SE_SE853  VKQYDQISIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
A_SE_SE889  VKQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
A_SE_UGSE8  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
A_UG_92UG0  VKQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
A_UG_U455_  VRQYDQILIE ICGKKTIGTV LVGPTPVNII GRNMLTLIGC TLNFPISPIS
AC_IN_2130  VRQYDQISIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
AC_RW_92RW  VKQYDQILIE ICGKKAIGTV LVGPTSVNII GRNMLTQLGC TLNFPISPIE
AC_SE_SE94  VKQYDQILIE ICGKKAIGTV LVGPTSVNII GRNMLTQIGC TLNFPISPIE
ACD_SE_SE8  VRQYDQILVE ICGHKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
ACG_BE_VI1  VRQYDQIMIE ICGKKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
AD_SE_SE69  VRQYDQILIE ICGYKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
AD_SE_SE71  VKQYDQILIE ICGKKAIGTV LVGPTPVNII GRNLLTQIGC ALNFPISPIE
ADHK_NO_97  VRQYDXILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
ADK_CD_MAL  VRQYDQILIE ICGKKAIGTI LVGPTPVNII GRNMLTQIGC TLNFPISPIE
AG_BE_VI11  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
AG_NG_92NG  VKQYDQILIE IEGKKAIGTV LVGPTPINII GRNMLTQIGC TLNFPISPIE
AGHU_GA_VI  VRQYDQVPIE ICGKKAIGTV LVGPTPINII GRNMLTQIGC TLNFPISPIE
AGU_CD_Z32  VRQYDQILIE IGEKRAIGTV LVGPTPINII GRNMLTQIGC TLNFPISPIE
AJ_BW_BW21  VRQYNDIHIE VEGKKAVGTV LIGPTPINII GRNILTQIGC TLNFPISPIE
B_AU_VH_AF  VRQYDQVLVE ICGHKAIGTV LIGPTPINII GRNMLTQLGC TLNFPISPIT
B_CN_RL42_  VRQYDQIPIE ICGHKAIGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIE
B_DE_D31_U  VRQYDQILIE ICGHKAIGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIK
B_DE_HAN_U  VRQYDQILVE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_FR_HXB2_  VRQYDQILIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_GA_OYI___ VRQYDQILIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_GB_CAM1_  VRQYDQIPIE ICGHKAIGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIE
B_GB_GB8_A  VKQYDQILVE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_GB_MANC_  VRQYDQILIE ICGHKAIGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIE
B_KR_WK_AF  VRQYDQVAIE ICGHKAIGTV LIGPTPVNII GRNLLTQIGC TLNFPISPIE
B_NL_3202A  VRQYDQIPIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_TW_TWCYS  VRQYDQIPIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPID
B_US_BC_L0  VRQYDQIPIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_US_DH123  VRQYDQVLIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_US_JRCSF  VRQYDQIPID ICGHKAVGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_US_MNCG_  VRQYDQITIG ICGHKAIGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIE
B_US_P896_  VRQYEQIDIE ICGHKAKGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_US_RF_M1  VRQYDQILIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_US_SF2_K  VRQYDQIPVE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_US_WEAU1  VRQYDQVPIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_US_WR27_  VRQYDQIPIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
B_US_YU2_M  VRQYDQIPIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
BF1_BR_93B  VRQYDQIPIE ICGRKATGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
C_BR_92BR0  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
C_BW_96BW0  VRQYEQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
C_BW_96BW1  VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIE
C_BW_96BW1  VRQYDQIVIE ICGKKAIGSV LVGPTPVNII GRNMLTQLGC TLNFPISPIK
C_BW_96BW1  VRQYDQILIE ICGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
C_ET_ETH22  VRQYDQIIIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGR TLNFPISPIE
C_IN_93IN1  VRQYDQIPIE ICGKKAIGTV LVGPTPINII GRNMLTQLGC TLNFPISPIE
C_IN_93IN9  VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
C_IN_93IN9  VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNLLTQLGC TLNFPISPIE
```

```
C_IN_94IN1  VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
C_IN_95IN2  VRQYEEIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
CRF01_AE_C  VRQYDQILIE ICGKRAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPID
CRF01_AE_C  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPID
CRF01_AE_C  VRQYDQIIIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPID
CRF01_AE_T  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPID
CRF01_AE_T  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPIS.PD
CRF01_AE_T  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPID
CRF01_AE_T  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPID
CRF01_AE_T  VKQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPID
CRF01_AE_T  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPID
CRF02_AG_F  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
CRF02_AG_F  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
CRF02_AG_G  VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
CRF02_AG_N  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
CRF02_AG_S  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
CRF02_AG_S  LRQYDQILIE ICGKKAMGSV LVGPTPVNII GKNILTQIGC TLNFPISPIE
CRF03_AB_R  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
CRF03_AB_R  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
CRF04_cpx_  VRQYDQIPIE ICGKKAIGTV LVGPTPVNII GRNMLTQLGC TLNFPISPIE
CRF04_cpx_  VKQYDQILIE ICGKKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
CRF04_cpx_  VRQYDQITIE ICGKKATGTV LVGPTPANII GRNMLTQLGC TLNFPISPIE
CRF05_DF_B  VRQYDQILVE ICGQKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
CRF05_DF_B  VRQYDQILIE ICGHKAVGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
CRF06_cpx_  VKQYDQILIE ICGKKAIGTV LVGPTPINII GRNMLTQIGC TLNFPISPIE
CRF06_cpx_  VRQYDQILIE ICGKRAMGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
CRF06_cpx_  VRQYDQILIE ICGKKAMGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
CRF06_cpx_  VRQYDQIPIE ICGKRAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
CRF11_cpx_  VKQYEEIIIE IEGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPID
CRF11_cpx_  VKQYEDITIE IEGKKAIGTV LIGPTPVNII GRNMLTQIGC TLNFPISPVD
D_CD_84ZR0  VRQYDHILIE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
D_CD_ELI_K  VRQYDQIPIE ICGQKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
D_CD_NDK_M  VRQYDQILIE ICGYKAMGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
D_UG_94UG1  VRQYDQIPLE ICGHKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
F1_BE_VI85  VKQYDNILIE ICGHKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPVSPIE
F1_BR_93BR  VKQYDSILIE ICGHRAIGTV LVGPTPVNII GRNMLTQIGC TLHFPISPIE
F1_FI_FIN9  VKQYDHILIE ICGHKAIGTV LVGPTPVNIV GRNMLTQIGC TLNFPISPIE
F1_FR_MP41  VKQYDQITID ICGHKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
F2_CM_MP25  VRQYDQVSIE ICGQKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
F2KU_BE_VI  VRQYDQVVME ICGQKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
G_BE_DRCBL  VRQYDQILIE ISGKRAIGTV LVGPTPINII GRNMLTQIGC TLNFPISPIE
G_NG_92NG0  VRQYDQILIE IGGKKAIGTV LVGPTPINII GRNMLTQIGC TLNFPISPIE
G_SE_SE616  VRQYDQVPIE ISGKKAIGTI LVGPTPINII GRNMLTQIGC TLNFPISPIE
H_BE_VI991  VRQYEQVAIE IFGKKAIGTV LVGPTPVNII GRNILTQMGC TLNLPISPIE
H_BE_VI997  VRQYDQVAIE ICGKKAIGTV LVGPTPVNII GRNILTQIGC TLNFPISPIE
H_CF_90CF0  VRQYEQVAIE ICGKKAIGTV LVGPTPVNII GRNILTQIGC TLNFPISPIE
J_SE_SE702  VRQYNEVPIE IEGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
J_SE_SE788  VRQYNEVPIE IEGKKAIGTV LIGPTPVNII GRNMLTQLGC TLNFPISPIE
K_CD_EQTB1  VRQYDQVCME ICGQKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE
K_CM_MP535  VRQYDQVLIE ICGQKAIGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
N_CM_YBF30  VRQYDNITVD IQGRKAVGTV LVGPTPVNII GRNLLTQIGC TLNFPISPIE
O_CM_ANT70  VKEYDNVTVE IEGREVQGTV LVGPTPVNII GRNILTGLGC TLNFPISPIA
O_CM_MVP51  VKEYNNVTVE VQGKEVQGTV LVGPTPVNIL GRNILTGLGC TLNFPISPIA
O_SN_99SE_  VKEYNQVPVE IEGREVLGTV LVGPTPVNII GRNILTGLGC TLNFPISPIA
O_SN_99SE_  VKEYNQVPVE IEGREVLGTV LVGPTPVNII GRNILTGLGC TLNFPISPIA
U_CD___83C  VRQYDQILIE ICGKKAIGTV LVGPTPVNII GRNMLTQIGC TLNFPISPIE

            201                                                 250
00BW0762_1  TVPVKLKPGM DGPKVRQWPL TEEKIKALTA ICDEMEKEGK ITKIGPENPY
00BW0768_2  TVPVKLKPGM NGPKVKQWPL TEEKIKALTA ICEEMEREGK ITKIGPENPY
00BW0874_2  TVPVKLKPGM DGPKVKQWPL TEEKIQALTA ICEEMEKEGK ITKIGPENPY
```

```
00BW1471_2  TVPVKLKPGM DGPKVKQWPL TEEKIKALKA ICEEMEKEGK ITKIGPKNPY
00BW1616_2  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW1686_8  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW1759_3  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW1773_2  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW1783_5  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW1795_6  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEDMEKEGK ITKIGPENPY
00BW1811_3  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW1859_5  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW1880_2  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW1921_1  TVPVKLKPGM DGPKVKQWPL TEEKIKALKE ICTEMEKEGK ITKIGPENPY
00BW2036_1  TVPVRLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW2063_6  TVPVKLKPGM DGPKVKQWPL TEEKIKALTG ICEEMEKEGK ITKIGPENPY
00BW2087_2  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW2127_2  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW2128_3  TVPLKLKPGM DGPKVNQWPL TEEKIKALTA ICEEMEKEGK IEKIGPENPY
00BW2276_7  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW3819_3  TVPVKLKPGM DGPKVKQWPL TEEKIKALME ICAEMEKEGK ITKIGPDNPY
00BW3842_8  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEKEGK IEKIGPENPY
00BW3871_3  TVPVKLKPGM DGPKVKQRPL TEEKIKALTA ICEDMEREGK ISKIGPENPY
00BW3876_9  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW3886_8  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW3891_6  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW3970_2  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
00BW5031_1  TVPVKLKPGM DGPKVKQWPL TKEKIEALTA ICEEMEKEGK ITKIGPENPY
 96BW01B21  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
  96BW0407  TVPVKLKPGM DGPKVKQWPL AEEKIKALTA ICEEMEKEGK ITKIGPENPY
  96BW0502  TVPVKLKPGM DGPKVKQWTL TEEKIKALTA ICEDMEKEGK ITKIGPENPY
 96BW06_J4  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICEEMEKEGK ITKIGPENPY
 96BW11_06  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICAEMEKEGK ITKIGPENPY
  96BW1210  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
 96BW15B03  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICEEMEKEGK VTKIGPENPY
 96BW16_26  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
 96BW17A09  TIPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
 96BWMO1_5  TVPVKLKPGM DGPKVKQWPL TEEKIKALKA ICEEMEKEGK ITKIGPENPY
 96BWMO3_2  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
98BWMC12_2  TVPVKLKPGM DGPKVKQWPL TAEKIKALTE ICEEMEKEGK ITKIGPENPY
98BWMC13_4  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
98BWMC14_a  TIPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
98BWMO14_1  TVPVKLKPGM DGPKVKQWPL TEEKIRALTA ICDEMEKEGK ITKIGPENPY
98BWMO18_d  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICEEMEKEGK ITKIGPENPY
98BWMO36_a  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
98BWMO37_d  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPH
99BW3932_1  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEKEGK IEKIGPENPY
99BW4642_4  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
99BW4745_8  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICEEMEKEGK ITKIGPENPY
99BW4754_7  TVPXKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ISKIGPENPY
99BWMC16_8  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEEEGK IEKIGPENPY
A2_CD_97CD  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEKEGK ISKIGPENPY
A2_CY_94CY  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEKEGK ISKIGPENPY
A2D___97KR  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEKEGK ISKIGPENPY
A2G_CD_97C  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEKEGK ISKIGPENPY
A_BY_97BL0  TVPVTLKPGM DGPKVKQWPL TEEKIKALXD ICKEXEKEGK ISKIXPENPY
A_KE_Q23_A  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
A_SE_SE659  TVPVKLKPGM DGPRVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
A_SE_SE725  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
A_SE_SE753  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
A_SE_SE853  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
A_SE_SE889  TVPVTLKPGM DGPRIKQWPL TEEKIKALTE ICREMEKEGK ISKIGPENPY
A_SE_UGSE8  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
A_UG_92UG0  TVPVKLKPGM DGPRIKQWPL TEEKIKALTE ICADMEREGR ISKIGPENPY
```

```
A_UG_U455_  TVPVKLKPEM DGPKVKQWPL TEEKIKALTE ICNEMEKEGK ISKIGPENPY
AC_IN_2130  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICDEMEKEGK ISKIGPENPY
AC_RW_92RW  TVPVALKPGM DGPKVKQWPL TEEKIKALRE ICTEMEKEGK ISKIGPENPY
AC_SE_SE94  TVPVTLKPGM DGPKVKQWPL TEEKIKALTE ICTDMEKEGK LSRIGPENPY
ACD_SE_SE8  TVPVKLKPGM DGPRVKQWPL TEEKIKALIE ICTEMEKEGK ISRIGPENPY
ACG_BE_VI1  TVPVKLKPGM DGPKVKQWPL TEEKIKALTD ICKEMEKEGK ISKIGPENPY
AD_SE_SE69  TVPVQLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK IARIGPENPY
AD_SE_SE71  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
ADHK_NO_97  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICLEMEKEGK ISKIGPENPY
ADK_CD_MAL  TVPVKLKPGM DGPRVKQWPL TEEKIKALTE ICKDMEKEGK ILKIGPENPY
AG_BE_VI11  TVPVKLKPGM DGPRVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
AG_NG_92NG  TVPVKLKPGI DGPKVKQWPL TEEKIKALTE ICTDMEKEGK ISKIGPENPY
AGHU_GA_VI  TVPVKLKPGI DGPKVKQWPL TEEKIKALTE ICNEMEQEGK ISRIGPENPY
AGU_CD_Z32  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
AJ_BW_BW21  TVPVNLKPGM DGPRVRQWPL TEEKIKALTE IFTEMEKEGK ISKIGPENPY
B_AU_VH_AF  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_CN_RL42_  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_DE_D31_U  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_DE_HAN_U  TVPVKLKPGM DGPKVKQWPL TEEKIKALIE ICTEMEKEGK ISKIGPENPY
B_FR_HXB2_  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_GA_OYI__  TVPVKLKPGM DGPKVKQWPL TEEKIKVLIE ICTEMEKEGK ISKVGPENPY
B_GB_CAM1_  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_GB_GB8_A  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_GB_MANC_  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_KR_WK_AF  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_NL_3202A  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_TW_TWCYS  TVPVKLKPGM DGPKVKQWPL TEEKIKVLIE ICTEMEKEGK ISKIGPENPY
B_US_BC_L0  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_US_DH123  TVPVKLKPGM DGPRVKQWPL SEEKIKALTE ICTEMEKEGK ISKIGPENPY
B_US_JRCSF  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_US_MNCG_  TVPVKLKPGM DGPKVKQWPL TEEKIKALIE ICTEMEKEGK ISKIGPENPY
B_US_P896_  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_US_RF_M1  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_US_SF2_K  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_US_WEAU1  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
B_US_WR27_  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK VTKIGPENPY
B_US_YU2_M  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
BF1_BR_93B  TVPVKLKPGM DGPRVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
C_BR_92BR0  TVPVKLKPGM DGPKVKQWLL TEEKIKALTA ICDEMEREGK ITKIGPENPY
C_BW_96BW0  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
C_BW_96BW1  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
C_BW_96BW1  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICEEMEKEGK VTKIGPENPY
C_BW_96BW1  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEKEGK ITKIGPENPY
C_ET_ETH22  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICEEMEQEGK ISRIGPENPY
C_IN_93IN1  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICDEMEKEGK ITKIGPENPY
C_IN_93IN9  TVPVKLKPGM DGPKVKQWPL TKEKIEALTA ICDEMEKEGK ITKIGPENPY
C_IN_93IN9  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICDEMEREGK ITKIGPENPY
C_IN_94IN1  TVPVKLKPGM DGPKVKQWPL TKEKIEALTI ICNEMEKEGK ITKIGPENPY
C_IN_95IN2  TVPVKLKPGM DGPKVKQWPL TEEKIKALTA ICDEMEKEGK ITKIGPENPY
CRF01_AE_C  TVPVTLKPGM DGPKVKQWPT ..EEKIALTE ICKEMEEEGK ISKIGPENPY
CRF01_AE_C  TVPVTLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEEEGK ISKIGPENPY
CRF01_AE_C  TVPVTLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEEEGK ISKIGPENPY
CRF01_AE_T  TVPVTLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEEEGK ISKIGPENPY
CRF01_AE_T  TVPVTLKPGM DGPKVKQWPL TEEKIKALTE TCKEMEEEGK ISKIGPENPY
CRF01_AE_T  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEEEGK ISKIGPENPY
CRF01_AE_T  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEEEGK ISKIGPENPY
CRF01_AE_T  TVPVTLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEEEGK ISKIGPENPY
CRF01_AE_T  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKKMEEEGK ISKIGPENPY
CRF02_AG_F  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
CRF02_AG_F  TVPVKLKPGM DGPKVKQWPL TEEKIKALTD ICAEMEKEGK ISKIGPENPY
CRF02_AG_G  TVPVKLKPGM DGPKVKQWPL TEEKIKALTD ICMEMEKEGK ISKIGPENPY
```

```
CRF02_AG_N  TVPVKLKPGM DGPKVKQWPL TEEKIKALTD ICTEMEKEGK ISKIGPENPY
CRF02_AG_S  TVPVKLKPGM DGPKVKQWPL TEEKIKALTD ICTEMEKEGK ISRIGPENPY
CRF02_AG_S  TVPLKLKPGM DGPKVKQWPL TEEKLKHLTD ICAEMEKEGK ISKIGPENPY
CRF03_AB_R  TVPVTLKPGM DGPKVKQWPL TEEKIKALTD ICKEMEKEGK ISKIGPENPY
CRF03_AB_R  TVPVTLKPGM DGPKVKQWPL TEEKIKALID ICKEMEKEGK ISKIGPENPY
CRF04_cpx_  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTDMEKEGK ISKIGPENPY
CRF04_cpx_  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
CRF04_cpx_  TVPVKLKPGM DGPKVKQWPL TEEKIKALRE ICTEMEQEGK ISKVGPENPY
CRF05_DF_B  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTDMEKEGK ISRIGPENPY
CRF05_DF_B  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTDMEKEGK ISRIGPENPY
CRF06_cpx_  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICKEMEKEGK ISKIGPENPY
CRF06_cpx_  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISRIGPENPY
CRF06_cpx_  TVPVKLKPGM DGPKIKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
CRF06_cpx_  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
CRF11_cpx_  TVPVQLKAGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
CRF11_cpx_  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
D_CD_84ZR0  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTDMEKEGK ISRIGPENPY
D_CD_ELI_K  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTDMEKEGK ISRIGPENPY
D_CD_NDK_M  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISRIGPENPY
D_UG_94UG1  TVPVKLKPGM DGPKVKQWPL TEEKIKALIE ICSELEKEGK ISKIGPENPY
F1_BE_VI85  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICLEMEKEGK ISKIGPENPY
F1_BR_93BR  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICMEMEKEGK ISKIGPENPY
F1_FI_FIN9  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTDMEKEGK ISRIGPENPY
F1_FR_MP41  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTDMEKEGK ISKIGPENPY
F2_CM_MP25  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
F2KU_BE_VI  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICLEMEKEGK ISKIGPENPY
G_BE_DRCBL  TVPVKLKPGM DGPRVKQWPL TEEKIKALTE ICNEMEKEGK ISKIGPENPY
G_NG_92NG0  TVPVKLKPGM DGPRVKQWPL TEEKIKALTE ICKDMEKEGK ISKIGPENPY
G_SE_SE616  TVPVKLKPGM DGPRVKQWPL TEEKIKALTE ICKEMEEEGK ISKIGPENPY
H_BE_VI991  TVPVTLKPGM DGPKVKQWPL TEEKIKALTE ICLEMEKEGK ISKIGPENPY
H_BE_VI997  TVPVKLKPGM DGPRVKQWPL TEEKIKALTE ICMEMEKEGK ISKIGPENPY
H_CF_90CF0  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISRIGPENPY
J_SE_SE702  TVPVKLKPGM DGPKIKQWPL TEEKIKALTQ ICAELEEEGK ISRIGPENPY
J_SE_SE788  TVPVKLKPGM DGPKIKQWPL TEEKIKALTQ ICAEMEEEGK ISRVGPENPY
K_CD_EQTB1  TVPVKLKPGM DGPKVKQWPL TEEKIKALVE ICTEMEKEGK ISKIGPENPY
K_CM_MP535  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICTEMEKEGK ISKIGPENPY
N_CM_YBF30  TVPVKLKPGM DGPKVKQWPL TTEKIEALRE ICTEMEKEGK ISRIGPENPY
O_CM_ANT70  PVPVKLKPGM DGPKVKQWPL SKEKIEALTA ICQEMEQEGK ISRIGPENPY
O_CM_MVP51  PVPVKLKPGM DGPKVKQWPL SREKIEALTA ICQEMEQEGK ISRIGPENPY
O_SN_99SE_  PVPVKLKPGM DGPKIKQWPL SKEKIEALTA ICQEMEQEGK ISRIGPENPY
O_SN_99SE_  PVPVKLKPGM DGPKIKQWPL SKEKIEALTA ICQEMEQEGK ISRIGPENPY
U_CD___83C  TVPVKLKPGM DGPKVKQWPL TEEKIKALTE ICLEMEKEGK ISKIGPENPY

            251                                                 300
00BW0762_1  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW0768_2  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW0874_2  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1471_2  NTPIFAIKKK DSTKWRKLVD FKELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1616_2  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1686_8  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1759_3  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1773_2  NTPVFAIKKK DSTKWRKLVD FGELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1783_5  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSL
00BW1795_6  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1811_3  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1859_5  NTPVFAIKKK DSTKWKKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1880_2  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW1921_1  NTPIFAIKKK DSTKWRKLVD FGELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW2036_1  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW2063_6  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW2087_2  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
```

```
00BW2127_2   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW2128_3   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWKVQLGIPH PAGLKKKKSV
00BW2276_7   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW3819_3   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKQKKSV
00BW3842_8   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW3871_3   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW3876_9   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW3886_8   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW3891_6   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW3970_2   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
00BW5031_1   NTPVSAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PVGLKKKKSV
 96BW01B21   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
  96BW0407   NTPVFAIKKK DSTKWRKLLD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
  96BW0502   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
 96BW06_J4   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
 96BW11_06   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
  96BW1210   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
 96BW15B03   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
 96BW16_26   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
 96BW17A09   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGSKKKKSV
 96BWMO1_5   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
 96BWMO3_2   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
98BWMC12_2   NTSVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
98BWMC13_4   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
98BWMC14_a   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
98BWMO14_1   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
98BWMO18_d   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
98BWMO36_a   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
98BWMO37_d   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
99BW3932_1   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGXKKKKSV
99BW4642_4   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
99BW4745_8   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
99BW4754_7   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
99BWMC16_8   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
A2_CD_97CD   NTPVFAIKKK DSDKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
A2_CY_94CY   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKAV
A2D___97KR   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKMKKSV
A2G_CD_97C   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
A_BY_97BL0   NTPVFAIKKK DSTKWRKLVX FXELNKRTQD FXEVQLGIPH PAGLKKKKSV
A_KE_Q23_A   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
A_SE_SE659   NTPIFAIKKK NSTRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKNSV
A_SE_SE725   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
A_SE_SE753   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
A_SE_SE853   NTPIFAIKKK DSNRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
A_SE_SE889   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FW.....IPH PAGKKKK.SV
A_SE_UGSE8   NTPIFAIKKK NSDRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
A_UG_92UG0   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
A_UG_U455_   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH TAGLKKKKSV
AC_IN_2130   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
AC_RW_92RW   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
AC_SE_SE94   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
ACD_SE_SE8   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
ACG_BE_VI1   NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKRSV
AD_SE_SE69   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
AD_SE_SE71   NTPIFAIKKK NSTRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
ADHK_NO_97   NTPVFAIKKK DSTKWXXXXD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
ADK_CD_MAL   NTPVFAIKKK DSTKWRKLVN FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
AG_BE_VI11   NTPIFAIKKK GSNRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKRKSV
AG_NG_92NG   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKRSV
AGHU_GA_VI   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FCEVQLGIPH PAGLKKKKSV
AGU_CD_Z32   NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVRLGIPH PARLKKKRSV
```

```
AJ_BW_BW21  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_AU_VH_AF  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_CN_RL42_  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEAQLGIPH PAGLKKKKSV
B_DE_D31_U  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_DE_HAN_U  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_FR_HXB2_  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_GA_OYI__  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_GB_CAM1_  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_GB_GB8_A  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PSGLKKKKSV
B_GB_MANC_  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_KR_WK_AF  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_NL_3202A  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSG
B_TW_TWCYS  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_US_BC_L0  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_US_DH123  NTPIFAIKKK NSTRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_US_JRCSF  NTPVFAIKKK DSTKWRKLVD FRELNRRTQD FWEVQLGIPH PAGLKKKKSV
B_US_MNCG_  NTPVFAIKKK DSTKWRKLVD FRELNKKTQD FWEVQLGIPH PAGLKKKKSV
B_US_P896_  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_US_RF_M1  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_US_SF2_K  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_US_WEAU1  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PSGLKKKKSV
B_US_WR27_  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
B_US_YU2_M  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
BF1_BR_93B  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
C_BR_92BR0  NTPVFAIKKK DSTKWRKLVD FRELNKRT.D FWEVQLGIPH PAGLKKKKSV
C_BW_96BW0  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
C_BW_96BW1  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
C_BW_96BW1  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
C_BW_96BW1  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
C_ET_ETH22  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
C_IN_93IN1  NTPIFAIKKK DSIKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
C_IN_93IN9  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEIQLGIPH PAGLKKKKSV
C_IN_93IN9  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
C_IN_94IN1  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
C_IN_95IN2  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF01_AE_C  NTPVFAIKEK DSTKWRKLVD FRELNKRTQD FWEVQLGVPH PAGLKKKKSV
CRF01_AE_C  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF01_AE_C  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF01_AE_T  NTPVFAIKKK DSTKWRKLVD FRELNKGTQD FWEVQLGIPH PAGLKKKKSV
CRF01_AE_T  NTPVFAIKKK D.STKRKLVG FRELNKRTQD FWEVQLGIPR PAGLKKKKSV
CRF01_AE_T  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF01_AE_T  NTPIFAIKKK NSTRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF01_AE_T  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF01_AE_T  NTPVFAIKKK DSTRWRKLVD FRELNKRTQD FWEVQLGIPR PAGLKKKKSV
CRF02_AG_F  NTPVFGIKKR DSTKWRKLVD FRELNKRTQD SWEVQLGIPH PAGLKKKKSV
CRF02_AG_F  NTPVFAIKKK DSTKWRKLVD FRELNKRTQG FWEVQLRIPH PAGLKKKKSV
CRF02_AG_G  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF02_AG_N  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF02_AG_S  NTPIFAIKKK NSTRWRKLVD FRELNKRTQD FWEIQLGIPH PAGLKKKKSV
CRF02_AG_S  NTPVFAIKRK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF03_AB_R  NTPVFAIKKK DSTKWRKLVG FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF03_AB_R  NTPVFVIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGLPH PAGLKKKKSV
CRF04_cpx_  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF04_cpx_  NTPIFAIKKK NSTRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF04_cpx_  NTPIFAIKKK NSNRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF05_DF_B  NTPIFAIKKK DSTRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSI
CRF05_DF_B  NTPIFAIKKK DSTRWRKLVN FRELNKKTQD FWEVQLGIPH PAGLKKKKSV
CRF06_cpx_  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF06_cpx_  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF06_cpx_  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKRKSV
CRF06_cpx_  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
```

```
CRF11_cpx_  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
CRF11_cpx_  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEIQLGIPH PAGLKKKKSV
D_CD_84ZR0  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSI
D_CD_ELI_K  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
D_CD_NDK_M  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
D_UG_94UG1  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
F1_BE_VI85  NTPVFAIKKK DSSKWRKLVD FKELNKRTQD FWEVQLGIPH PAGLKKKKSV
F1_BR_93BR  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
F1_FI_FIN9  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
F1_FR_MP41  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
F2_CM_MP25  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKRSV
F2KU_BE_VI  NTPIFAIKKK NSNRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
G_BE_DRCBL  NTPIFAIKKK DSTRWRKLVD FRELNKRTQD FWEVPLGIPH PGGLKQKRSV
G_NG_92NG0  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKRSV
G_SE_SE616  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
H_BE_VI991  NTPIFAIKKK NSTRWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
H_BE_VI997  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
H_CF_90CF0  STPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
J_SE_SE702  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
J_SE_SE788  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
K_CD_EQTB1  NTPVFAIKKK DSTKWIKLVD FRELNKRTPD FWEVQLGIPH PAGLKKKKSV
K_CM_MP535  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV
N_CM_YBF30  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKQKKSV
O_CM_ANT70  NTPIFAIKKK DGTKWRKLVD FRELNKRTQE FWEVQLGIPH PGGLKQKQSV
O_CM_MVP51  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PGGLKQRQSV
O_SN_99SE_  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PGGLKQKQSV
O_SN_99SE_  NTPIFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PGGLKQKQSV
U_CD___83C  NTPVFAIKKK DSTKWRKLVD FRELNKRTQD FWEVQLGIPH PAGLKKKKSV

            301                                                 350
00BW0762_1  TVLDVGDAYF SVPLDENFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW0768_2  TVLDVGDAYF SVPLDENFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW0874_2  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLSQGWKGS
00BW1471_2  TVLDVGDAYF SVPLDRDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW1616_2  TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW1686_8  TVLDVGDAYF SVPLDESFRK YTAFTIPSRN NETPGIRYQY NVLPQGWKGS
00BW1759_3  TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGVRYQY NVLPQGWKGS
00BW1773_2  TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NEAPGIRYQY NVLPQGWKGS
00BW1783_5  TVLDVGDAYF SVPLHENFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW1795_6  TVLDVGDAYF SIPLDEGFRK YTAFTIPSIN NATPGIRYQY NVLPQGWKGS
00BW1811_3  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW1859_5  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW1880_2  TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW1921_1  TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW2036_1  TVLDVGDAYF SVPLDENFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW2063_6  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW2087_2  TVLDVGDAYF SVPLDESFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
00BW2127_2  TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW2128_3  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW2276_7  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW3819_3  TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW3842_8  TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW3871_3  TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW3876_9  TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW3886_8  TVLDVGDAYF SVPLDESFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
00BW3891_6  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW3970_2  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
00BW5031_1  TVLDVGDAYF SVPLDENFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
 96BW01B21  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
  96BW0407  AVLDVGDAYF SVPLDESFRK YTAFTIPSIN NSTPGIRYQY NVLPQGWNGS
  96BW0502  TVLDMGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
```

```
96BW06_J4   TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGITYQY NVLPQGWKGS
96BW11_06   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
 96BW1210   TVLDVGDAYF SVPLDESFRK YTAFTIPSRN NETPGIRYQY NVLPQGWKGS
96BW15B03   TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
96BW16_26   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NATPGIRYQY NVLPQGWKGS
96BW17A09   TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
96BWMO1_5   TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
96BWMO3_2   TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
98BWMC12_2  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
98BWMC13_4  TVLDVGDAYF SVPLDEGFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
98BWMC14_a  TVLDVGDAYF SVPLDEDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
98BWMO14_1  TVLDVGDAYF SVPLDESFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
98BWMO18_d  TVLDVGDAYF SVPLDENFRK YTAFTIPSIN NATPGIRHQY NVLPQGWKGS
98BWMO36_a  TVLDVGDAYF SVPLDEDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
98BWMO37_d  TVLDVGDAYF SVPLDENFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
99BW3932_1  TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NATPGIRYQY NVLPQGWKGS
99BW4642_4  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
99BW4745_8  TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
99BW4754_7  TVLDVGDAYF SIPLDENFRK YTAFTIPSTN NATPGVRYQY NVLPQGWKGS
99BWMC16_8  TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
A2_CD_97CD  TVLDVGDAYF SVPLHEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
A2_CY_94CY  TVLDVGDAYF SVPLHEDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
A2D___97KR  TVLDVRDAYF SVPLHEDFRK YTAFTIPSTN NETPGVRYQY NVLPQGWKGS
A2G_CD_97C  TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGVRYQY NVLPQGWKGS
A_BY_97BL0  TVLDVGDAYF SVPLDESFRK XXAFTIPSVN NETPGVRYQY NVLPQGWKGS
A_KE_Q23_A  TVLDVGDAYF SVPLHEEFRK YTAFTIPSTN NETPXIRYQY NVLPQGWKGS
A_SE_SE659  TVLDVGDAYF SVPLHEDFRK YTAFTIPSTN NETPGVRYQY NVLPQGWKGS
A_SE_SE725  TVLDVGDAYF SVPLDESFRK YTAFTIPSTN NATPGIRYQX NVLPQGWKGS
A_SE_SE753  TVLDLGDAYF SVPLHEGFRK YTAFTIPSTN NETPGVRYQY NVLPQGWKGS
A_SE_SE853  TVLDVGDAYF SVPLDESFRK YTAFTIPSTN NATPGIRYQY NVLPQGWKGS
A_SE_SE889  TVLDVGDAYF SVPLDKNFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
A_SE_UGSE8  TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
A_UG_92UG0  TVLDVGDAYF SVPLDESFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
A_UG_U455_  TVLDVGDAYF SVPLDESFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
AC_IN_2130  TVLDVGDAYF SVPLYEEFRK YTAFTIPSIN NETPGVRYQY NVLPQGWKGS
AC_RW_92RW  TVLDVGDAYF SVPLDESFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
AC_SE_SE94  TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
ACD_SE_SE8  TILDVGDAYF SVPLDKDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
ACG_BE_VI1  TVLDVGDAYF SVPLDESFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
AD_SE_SE69  TVLDVGDAYF SVPLDKEFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
AD_SE_SE71  TVLDVGDAYF SVPLYEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
ADHK_NO_97  TVLDVGDAYF SVPLHEEFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
ADK_CD_MAL  TVLDVGDAYF SVPLAEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
AG_BE_VI11  TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
AG_NG_92NG  TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
AGHU_GA_VI  TVLDVGDAYF SIPLDENFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
AGU_CD_Z32  TVLDVGDAYF SVPLYEGFRR YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
AJ_BW_BW21  TVLDVGDAYF SVPLHEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_AU_VH_AF  TVLDVGDAYF SVPLDENFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
B_CN_RL42_  TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_DE_D31_U  TVLDVGDAYF SVPLDKDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
B_DE_HAN_U  TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGVRYQY NVLPQGWKGS
B_FR_HXB2_  TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_GA_OYI___ TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_GB_CAM1_  TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_GB_GB8_A  TVLDVGDAYF SVPLDKDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
B_GB_MANC_  TVLDVGDAYF SVPLDKEFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_KR_WK_AF  TVLDVGDAYF SVPLYEDFRK YTVFTIPSIN NEAPGVRYQY NVLPQGWKGS
B_NL_3202A  TALDVGDAYF SVPLDEDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
B_TW_TWCYS  TVLDVGDAYF SVPLDKDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
B_US_BC_L0  TVLDVGDRYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
```

```
B_US_DH123   TVLDVGDAYF SIPLDEDFRK YTAFTIPSVN NAAPGIRYQY NVLPQGWKGS
B_US_JRCSF   TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_US_MNCG_   TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_US_P896_   TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_US_RF_M1   TVLDVGDAYF SVPLDKEFRK YTAFTIPSIN NETPRIRYQY NVLPQGWKGS
B_US_SF2_K   TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_US_WEAU1   TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
B_US_WR27_   TVLDVGDAXF SVXLDXEXRK YTAFTIPSXH NETPGIRYQY NVLPQGWKGS
B_US_YU2_M   TVLDVGDAYF SVPLHEDFRK YTAFTIPSIN NETPGTRYQY NVLPQGWKGS
BF1_BR_93B   TVLDVGDAYF SVPLDKDFRK YTAFTIPSTN NETPGLRYQY NVLPQGWKGS
C_BR_92BR0   TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
C_BW_96BW0   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NSTPGIRYQY NVLPQGWKGS
C_BW_96BW1   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
C_BW_96BW1   TVLDVGDAYF SVPLDESFRK YTAFTIPSRN NETPGIRYQY NVLPQGWKGS
C_BW_96BW1   TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
C_ET_ETH22   TVLDVGDAYF SVPLDEGFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
C_IN_93IN1   TVLDVGDAYF SVPLYEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
C_IN_93IN9   TVLDVGDAYF SVPLYEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
C_IN_93IN9   TVLDVGDAYF SVPLYEDFRK YTAFTIPSRN NETPGIRYQY NVLPQGWKGS
C_IN_94IN1   TVLDVGDAYF SVPLYEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
C_IN_95IN2   TVLDVGDAYF SVPLYEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF01_AE_C   TVLDVGDAYF SVPLYEGFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF01_AE_C   TVLDVGDAYF SVPLDEGFRK YTAFTIPSIN NETPGVRYQY NVLPQGWKGS
CRF01_AE_C   TVLDVGDAYF SVPLHE.SRK YTAFTIPSIN NETPGIRYQY NVLPQGWRGS
CRF01_AE_T   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN DETPGIRYQY NVLPQGWKGS
CRF01_AE_T   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF01_AE_T   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF01_AE_T   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF01_AE_T   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF01_AE_T   TVLDVGDAYF SVPLDESFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF02_AG_F   TVLDVGDAYF SVPLDKDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
CRF02_AG_F   TVLDVGDAYF SVPLDKDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
CRF02_AG_G   TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF02_AG_N   TVLDVGDAYF SVPLDKDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
CRF02_AG_S   TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF02_AG_S   TVLDVGDAYF SVPLDEDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
CRF03_AB_R   TVLDVGDAYF SVPLDQDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
CRF03_AB_R   TVLDVGDAYF SVPLDQDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
CRF04_cpx_   TVLDVGDAYF SVPLDPEFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
CRF04_cpx_   TVLDVGDAYF SVPLDPAFRK YTAFTIPSTN NETPGVRYQY NVLPQGWKGS
CRF04_cpx_   TVLDVGDAYF SVPLDPEFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF05_DF_B   TVLDVGDAYF SVPLDKEFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF05_DF_B   TVLDVGDAYF SVPLHEDFRK YTAFTIPSIN NETPGFRYQY NVLPQGWKGS
CRF06_cpx_   TVLDVGDAYF SIPLDEKFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
CRF06_cpx_   TVLDVGDAYF SIPLDKDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
CRF06_cpx_   TVLDVGDAYF SVPLGENFRK YTAFTIPSLN NETPGIRYQY NVLPQGWKGS
CRF06_cpx_   TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
CRF11_cpx_   TVLDVGDAYF SVPLDENFRK YTAFTIPSIN NETPGVRYQY NVLPQGWKGS
CRF11_cpx_   TVLDVGDAYF SVLLDESFRK YTAFTIPSLN NETPGIRYQY NVLPQGWKGS
D_CD_84ZR0   TVLDVGDAYF SIPLCEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
D_CD_ELI_K   TVLDVGDAYF SVPLDEDFRK YTAFTISSIN NETPGIRYQY NVLPQGWKGS
D_CD_NDK_M   TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
D_UG_94UG1   TVLDVGDAYF SVPLHEDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
F1_BE_VI85   TVLDVGDAYF SVPLDKDFKK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
F1_BR_93BR   TVLDVGDAYF SVPLDKDFRK YTASTIPSTN NETPGVRYQY NVLPQGWKGS
F1_FI_FIN9   TVLDVGDAYF SVPLDKDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
F1_FR_MP41   TVLDVGDAYF SVPLDKEFRK YTAFTIPSLN NETPGIRYQY NVLPQGWKGS
F2_CM_MP25   TVLDVGDAYF SVPLDKEFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
F2KU_BE_VI   TVLDVGDAYF SVPLDPEFRK YTAFTIPSVN NETPGVRYQY NVLPQGWKGS
G_BE_DRCBL   TVLDVGDAYF SVPLDENFRK YTAFTIPSTN NETPGIRYQY ...PQGWKGS
G_NG_92NG0   TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
```

```
G_SE_SE616   TVLDVGDAYF SVPLDEDFRK YTAFTIPSIN NETPGVRYQY NVLPQGWKGS
H_BE_VI991   SVLDVGGAYF SVPLHEDFRK YTAFTIPSTN NETPGIRYQY NVLPQGWKGS
H_BE_VI997   SVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
H_CF_90CF0   SVLDVGDAYF SVPLDKEFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
J_SE_SE702   TVLDVGDAYF SVPLYEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
J_SE_SE788   TVLDVGDAYF SVPLYEDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
K_CD_EQTB1   TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
K_CM_MP535   TVLDVGDAYF SVPLDKDFRK YTAFTIPSIN NETPGVRYQY NVLPQGWKGS
N_CM_YBF30   TVLDVGDAYF SCPLDKDFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS
O_CM_ANT70   TVLDVGDAYF SCPLDPDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
O_CM_MVP51   TVLDVGDAYF SCPLDPDFRK YTAFTIPSVN NETPGVRYQY NVLPQGWKGS
O_SN_99SE_   TVLDVGDAYF SCPLDPDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
O_SN_99SE_   TVLDVGDAYF SCPLDPDFRK YTAFTIPSVN NETPGIRYQY NVLPQGWKGS
U_CD___83C   TVLDVGDAYF SVPLDENFRK YTAFTIPSIN NETPGIRYQY NVLPQGWKGS

             351                                                 400
00BW0762_1   PAIFQSSMTR ILEPFRTQNP EIVIYQYMDD LYVGSDLEIG QHRVKIEELR
00BW0768_2   PAIFQDSMTK ILEPFRAQNP EIVIYQYMDD LYVGSNLEIG QHRAKIEELK
00BW0874_2   PAIFQSSMTR ILEPFRAQNP ELVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW1471_2   PAIFQSSMTK ILEPFRAQNP EIVIYQYMDN LYVRSDLEIG QHRAKIEELR
00BW1616_2   PAIFQSSMTK ILEPFRAKNP EIVIYQYMDD LYVGSDLEIG QHREKIEELR
00BW1686_8   PAIFQSTMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW1759_3   PAIFQSSMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW1773_2   PSIFQSSMTK ILEPFRTQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW1783_5   PAIFQSSMTK ILGPFRTQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEGLR
00BW1795_6   PAIFQSSMTR ILEPFRTQNP DIVIYQYMDD LYVGSDLEIG KHRAKIEELR
00BW1811_3   PAIFQSSMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEKLR
00BW1859_5   PAIFQSSMTR ILEPFRTQNP EIVIYQYMDD LYVESDLEIG QHRAKIEELR
00BW1880_2   PAIFQSSMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLKIG QHRAKIEELK
00BW1921_1   PAIFQSSMTK ILEPFRAQNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW2036_1   PAIFQSSMTK ILEPFRAKNP ELVIYQYMDD LYVGSDLEIG QHRARIEELR
00BW2063_6   PAIFQSSMTK ILEPFRAKNP DIVIYQYMDD LYVGSDLDIG QRREKIEDLR
00BW2087_2   PAIFQCSMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW2127_2   PAIFQSSMTK ILEPFRARNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW2128_3   PSIFQSSMTK ILKPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIQELR
00BW2276_7   PAIFQSSMTR ILEPFRAQNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW3819_3   PAIFQCSMTK ILEPFRAKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW3842_8   PAIFQSSMTK ILEPFRAKNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW3871_3   PAIFQSSMTR ILEPFRTQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW3876_9   PAVFQSSMTK ILEPFRVQNP EIVVYQYMDD LYVGSDLEIG QHRAKIEELR
00BW3886_8   PAIFQCSMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW3891_6   PAIFQSSMTR ILEPFRAQNP GIVIYQYMDD LYVGSNLEIG QHRAKIEELR
00BW3970_2   PAIFQSSMTX ILEPFRARNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
00BW5031_1   PAIFQSSMTK ILEPFRAKNP EIVIYQYMDD LYVGSDLEIV QHRAKVEELR
 96BW01B21   PAIFQSSMTK ILEPFRALNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
  96BW0407   PAIFQSSMTK ILEPFRTKNP DILIYQYMDD LYVGSDLQIG QHRAKLEELT
  96BW0502   PAIFQSSMTK ILEPFRLQNP EIVIYQYMDD LYVGSDLEIG QRRAQIEELR
 96BW06_J4   PAIFQSSMTR ILEPFRTQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
 96BW11_06   PSIFQSSMTK ILEPFRAKNP ELVIYQYMDD LYVGSDLEIG QHRAKIEELR
  96BW1210   PAIFQSSMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHTAKIEELR
 96BW15B03   PSIFQSSMTK ILEPFRARNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
 96BW16_26   PAIFQSSMTK ILEPFRAQNP GIVIYQYMDD LYVGSDLEIG QHRAKIEELR
 96BW17A09   PAIFQSSMTK ILEPFRAQNP EIVIYQYMDD LYVGXDLEIG QHRAKIEELR
 96BWMO1_5   PAIFQSSMTK ILELFRAKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
 96BWMO3_2   PAIFQASMIK ILEPFRAQNP EMVIYQYMDD LYVGSDLEIG QHRAKIEELR
98BWMC12_2   PAIFQSSMTR ILEPFRAQNP EIVIYQYMDD LYVGSDLDIG QHRAKIEELR
98BWMC13_4   PAIFQCSMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG KHRAKIEELR
98BWMC14_a   PAIFQSSMTK ILEPFRARNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
98BWMO14_1   PAIFQCSMTK ILEPFRAKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
98BWMO18_d   PAIFQSSMTK ILEPFRTQNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
98BWMO36_a   PAIFQGSMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
```

```
98BWMO37_d  PAIFQCSMTR ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
99BW3932_1  PAIFQSSMTK ILKPFREQNP EMVIYQYMDD LYVGSDLEIG QHTAKIEELR
99BW4642_4  PAIFQSSMTR ILEPFRIQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
99BW4745_8  PAIFQSSMTK ILEPFRAKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
99BW4754_7  PAIFQSSMTK ILEPFRAQNP DIVIYQYMDD LYVGSDLEIG QHRAKIEKLR
99BWMC16_8  PAIFQSSMTK ILEPFRTKNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
A2_CD_97CD  PAIFQSSMTK ILDPFRARNP EMVIYQYMDD LYVGSDLEID QHRAKIEELR
A2_CY_94CY  PAIFQSSMTK ILEPFRSKNT ELIIYQYMDD LYVGSDLEIS QHRVKIEELR
A2D___97KR  PAIFQSSMTK ILEPFRKQNP EIVIYRYMDD LYVGSDLEIG QHRTKIEELR
A2G_CD_97C  PAIFQSSMTK ILEPFRADNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
A_BY_97BL0  PSIFQSSMTK ILEPFRLKNP EIVIYQYXDD LYVGSDLETG QHRTKIEELR
A_KE_Q23_A  PAIFQSSMTK ILEPFRSKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
A_SE_SE659  PAIFQSSMTK ILEPFRSKNP DIIIYQYMDD LYVGSDLEIG QHRAKIEELR
A_SE_SE725  PAIFQSSMTK ILEPFRLKNP EIIIYQYMDD LYVGSDLEIG QHRTKIEELR
A_SE_SE753  PAIFQSSMTK ILEPFRERNP EVIIYQYMDD LYVGSDLEIG QHRTKIEELR
A_SE_SE853  PSIFQSSMTK ILEPFRSKNP EIIIYQYMDD LYVGSDLEIG QHRTKIEELR
A_SE_SE889  PAIFQSSMIK ILEPFRVQNP EIIIYQYMDD LYVGSDLEIG QHRAKVEELR
A_SE_UGSE8  PAIFQSSMTK ILEPFRSKNP EIIIYQYMDD LYVGSDLEIE QHRTKIEELR
A_UG_92UG0  PAIFQASMTK ILEPFRSKNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
A_UG_U455_  PSIFQSSMTK ILEPFRSQHP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
AC_IN_2130  PAIFQASMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLKIG QHRAKIEELR
AC_RW_92RW  PAIFQNSMTK ILEPFRAQNQ EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
AC_SE_SE94  PAIFQSSMTK ILAPFRSQNP EIIIYQYMDD LYVGSDLEIG QHRTKIEELR
ACD_SE_SE8  PAIFQSSMTK ILEPFRSKNP DMIIYQYMDD LYVGSDLEIG QHRTKIEELR
ACG_BE_VI1  PAIFQASMTK ILDPFRRENP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
AD_SE_SE69  PAIFQSSMTK ILEPFRKQNP EMVIYQYMDD LYVGSDLEIG QHRIKIEELR
AD_SE_SE71  PAIFQSSMTK ILEPFRSKNP ELIIYQYMDD LYVGSDLEIG QHRIKIEELR
ADHK_NO_97  PAIFQCSMTK ILEPFRAKNP EIVIYQYMDD LYVGSDLEIG QHRTKIEELR
ADK_CD_MAL  PAIFQSSMTK ILEPFRTKNP EIVIYQYMDD LYVGSDLEIG QHRTKIEELR
AG_BE_VI11  PAIFQASMTK ILEPFRTENP KIVIYQYMDD LYVGSDLEIG QHRAKIDELR
AG_NG_92NG  PAIFQSSMTK ILEPFRTENP EIVIYQYMDD LYVGSDLETG QHRAKIEELR
AGHU_GA_VI  PAIFQSSMTK ILEPFRKQNP EMVIYQYMDD LYVGSDLEIG QHRAKIEELR
AGU_CD_Z32  PAIFQSSMTK ILEPFRTKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
AJ_BW_BW21  PAIFQYSMTK ILEPFRANNP EIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_AU_VH_AF  PAIFQSSMTK ILEPFRKQNP DVVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_CN_RL42_  PAIFQCSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
B_DE_D31_U  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_DE_HAN_U  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIE QHRTKIEELR
B_FR_HXB2_  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_GA_OYI__  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_GB_CAM1_  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_GB_GB8_A  PAIFQSSMTK ILEPFRKQNP EIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_GB_MANC_  PAIFQSSMTK ILEPFRKQNP DVVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_KR_WK_AF  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRIKIEELR
B_NL_3202A  PAIFQCSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_TW_TWCYS  PAIFQSSMTR ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRIKVEELR
B_US_BC_L0  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_US_DH123  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIE QHRTKIEELR
B_US_JRCSF  PAIFQSSMTK ILEPFRKQNP DIIIYQYMDD LYVGSDLEIG QHRTKIEELR
B_US_MNCG_  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
B_US_P896_  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRAKIEDLR
B_US_RF_M1  PAIFQSSMTK ILEPFKKQNP EIVIYQYMDD LYVGSDLEIG QHRIKIEELR
B_US_SF2_K  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_US_WEAU1  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
B_US_WR27_  PTIFPSSMTQ ILEPFRKPNP EIVIYQYMDD LYVGSNLEIG QHRTKIEELR
B_US_YU2_M  PAIFQSSMTT ILEPFRKQNP DLVIYQYMDD LYVGSDLEIG QHRTKIEELR
BF1_BR_93B  PAIFQSSMTK ILEPFRKQNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
C_BR_92BR0  PSIFQSSTTK ILEPFRAQNP EIIIYQYMDD LYVGSDLEIG QHRAKIEELR
C_BW_96BW0  PAIFQSSMIK ILEPFRTKNP DIVIYQYMDD LYVGSDLEIG QHRAKIEELR
C_BW_96BW1  PSIFQSSMTK ILEPFRAKNP ELVIYQYMDD LYVGSDLEIG QHRARIEELR
C_BW_96BW1  PAIFQSSMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHTAKIEELR
```

```
C_BW_96BW1   PSIFQSSMTK ILEPFRARNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
C_ET_ETH22   PPIFQSSMPQ ILEPFRAPNP EIVIYQYMDD LYVGSDLEIG QHRAPIEELR
C_IN_93IN1   PAIFQSSMIR ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
C_IN_93IN9   PAIFQSSMTR ILEPFRARNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
C_IN_93IN9   PAIFQASMTK ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
C_IN_94IN1   PAIFQSSMTK ILEPFRGRNP EIDIYQYMDD LYVGSDLEIG QHRAKIEELR
C_IN_95IN2   PAIFQNSMTR ILEPFRAQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF01_AE_C   PAIFQSSMTK ILEPYRAKNP EIVIYQYMDD LYVGSDLEIG QHRTKIEELR
CRF01_AE_C   PAIFQSSMTK ILEPFRARNP EIVIYQYMDD LYVGSDLEIG QHRTKVEDLR
CRF01_AE_C   PAIFQCSMTK ILEPFRAKNP EIVIYQYMDD LYVGSDLEIG QHGTKIEELR
CRF01_AE_T   PAIFQSSMTK ILEPFRIKNP EMVIYQYMDD LYVGSDLEIG QHRTKIEELR
CRF01_AE_T   PAIFQSSMTK ILEPFRIKNP EMVIYQYMDD LYVGSDLEIG QHRIKIEELR
CRF01_AE_T   PAIFQSSMTK ILEPFRIKNP EMVIYQYMDD LYVGSDLEIG QHRTKIEELR
CRF01_AE_T   PAIFQSSMTK ILEPFRIKNP EMVIYQYMDD LYVGSDLEIG QHRTKIEELR
CRF01_AE_T   PAIFQSSMTK ILEPFRIKNP EMVIYQYKDD LYVGSDLEIG QHRTKIEELR
CRF01_AE_T   PAIFQCSTTK ILEPFRTKNP EIVIYQYMDD LYVGSDLEIG QHRTKIEELR
CRF02_AG_F   PAIFQASMTN ILEHYRIKNP EIMIYQYMDD LYVGSDLEIE QHRAKIEELR
CRF02_AG_F   PAIFQASMTK ILEPFRIKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF02_AG_G   PAIFQSSMTK ILEPYRIKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF02_AG_N   PAIFQASMTK ILEPFRTKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF02_AG_S   PAIFQASMTK ILEPFRTKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF02_AG_S   PAIFQASMTK ILEPFRTKNP ELVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF03_AB_R   PAIFQSSMTK ILEPFRKQNP EIVIYQYMDD LYVGSDLEIG QHRTEIEELR
CRF03_AB_R   PAIFQSSMTK ILEPFRKQNP EIVIYQYMDD LYVGSDLEIG QHRIKIEELR
CRF04_cpx_   PAIFQCSMTK ILEPFRFKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF04_cpx_   PAIFQYSMTK ILEPFRTRNP EIVIYQYMDD LYVGSDLEIG QHRTKIEELR
CRF04_cpx_   PAIFQCSMTK ILEPFRTKNP EMVIYQYMDD LYVGSDLEIG QHKAKIEELR
CRF05_DF_B   PAIFQCSMTK ILEPFRKQNP EVVIYQYMDD LYVGSDLEIG QHRAKIKELR
CRF05_DF_B   PAIFQCSMTK ILEPFRKQNP EMVIYQYMDD LYVGSDLEIG QHRTKIEELR
CRF06_cpx_   PAIFQSSMIK ILEPFRIKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF06_cpx_   PAIFQCSMTK ILDPFRIKNP ELVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF06_cpx_   PAIFQSSMIK ILEPFRTKNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
CRF06_cpx_   PAIFQSSMIK ILEPFRQKNP EIVIYQYMDD LYVGSDLEIG QHREKVEELR
CRF11_cpx_   PAIFQSSMTK ILEPFRTQNP EIVIYQYMDD LYVGSDLEIG QHREKVEELR
CRF11_cpx_   PAIFQSSMTK ILEPFRTQNP EVVIYQYMDD LYVGSDLEIG QHREKVEELR
D_CD_84ZR0   PAIFQSSMIK ILEPFRKQNP EVVIYQYMDD LYVGSDLEIG QHRAKIEKLR
D_CD_ELI_K   PAIFQSSMTK ILEPFRKQNP EMVIYQYMDD LYVGSDLEIG QHRTKIEKLR
D_CD_NDK_M   PAIFQSSMTK ILEPFRKQNP EIVIYQYMDD LYVGSDLEIG QHRTKIEELR
D_UG_94UG1   PAIFQSSMTK ILEPFRKQNP EMIIYQYMDD LYVGSDLEIG QHRIKIEELR
F1_BE_VI85   PAIFQCSMTK ILEPFRMKNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
F1_BR_93BR   PAIFQYSMTK ILDPFRAKNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
F1_FI_FIN9   PAIFQCSMTK ILEPFRTRNP DIVIYQYMDD LYVGSDLEIG QHRTKIEELR
F1_FR_MP41   PAIFQSSMTK ILEPFRAKNP DIVIYQYMDD LYVGSDLELG QHRMKIEELR
F2_CM_MP25   PAIFQSSMIK ILEPFRKENP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
F2KU_BE_VI   PAIFQYSMTK ILEPFRTKNP EMVIYQYMDD LYVGSDLEIG QHRTKIEELR
G_BE_DRCBL   PAIFQSSMTK ILEPFRTQNP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
G_NG_92NG0   PAIFQSSMTK ILEPSRTKNP EMVIYQYMDD LYVGSDLEIG QHRAKIEELR
G_SE_SE616   PAIFQSSMTR ILEPFRANNP EMVIYQYMDD LYVGSDLEIG QHRAKIEELR
H_BE_VI991   PAIFQSSMTK ILEPFRKQNP EVIIYQYMDD LYVGSDLEIG QHREKIEELR
H_BE_VI997   PAIFQSSMTK ILEPFRKQNP EIIIYQYMDD LYVGSDLEIG QHRAKIEELR
H_CF_90CF0   PAIFQSSMTK ILAPFREQNP EMVIYQYMDD LYVGSDLEIG QHRAKIEELR
J_SE_SE702   PAIFQCSMTK ILKPFRERNP EIVIYQYMDD LYVGSDLEIE QHRRKIKELR
J_SE_SE788   PAIFQCSMTK ILKPFRERNP EIVIYQYMDD LYVGSDLEIE QHRRKIKELR
K_CD_EQTB1   PAIFQCSMTK ILEPFRRKNP DMVLYQYMDD LYVGSDLEIG QHRAKIEELR
K_CM_MP535   PAIFQHSMTK ILEPFRIKNP EMVIYQYMDD LYVGSDLEIG QPRTKIEELR
N_CM_YBF30   PAIFQSTMTK ILEPFREKHP EIIIYQYMDD LYVGSDLELA QHREAVEDLR
O_CM_ANT70   PAIFQSSMTK ILDPFRRDNP ELEICQYMDD LYVGSDLPLT EHRKRIELLR
O_CM_MVP51   PAIFQSSMTK ILDPFRKSNP EVEIYQYIDD LYVGSDLPLA EHRKRVELLR
O_SN_99SE_   PAIFQSSMTK ILDPFRKDNP ELEICQYMDD LYVGSDLPLT EHRKRVELLR
O_SN_99SE_   PAIFQSSMTK ILDPFRKNNP ELEICQYMDD LYVGSDLPLT EHRKRVELLR
U_CD___83C   PAIFQSSMTK ILEPFRKENP EIVIYQYMDD LYVGSDLEIG QHRAKIEELR
```

```
            401                                                      450
00BW0762_1  RHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKEDWTVND
00BW0768_2  DHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKNSWTVND
00BW0874_2  AHLLKWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW1471_2  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW1616_2  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW1686_8  DHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQTIQL PTKESWTVND
00BW1759_3  NHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW1773_2  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIKL PDKDSWTVND
00BW1783_5  NHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
00BW1795_6  EHLLKWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
00BW1811_3  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKD.STVND
00BW1859_5  DHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW1880_2  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKEDWTVND
00BW1921_1  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIIL PEQESWTVND
00BW2036_1  KHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIKL PEKESWTVNG
00BW2063_6  NHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKESWTVND
00BW2087_2  GHLLQWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW2127_2  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKGTVQPIQL PEKDSWTVND
00BW2128_3  DHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW2276_7  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW3819_3  GHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW3842_8  AHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW3871_3  DHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIKL PEKESWTVND
00BW3876_9  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQKIQL PTKDSWTVND
00BW3886_8  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIKL PEKESWTVND
00BW3891_6  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
00BW3970_2  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWAVQPIQL PEKDSWTVND
00BW5031_1  KHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
 96BW01B21  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPINL PEKESWTVND
  96BW0407  EHLLIWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
  96BW0502  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
 96BW06_J4  NHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PTKDSWTVND
 96BW11_06  KHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQSIKL PEKESWTVND
  96BW1210  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKESWTVND
 96BW15B03  KHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIEL PEKESWTVND
 96BW16_26  THLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
 96BW17A09  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
 96BWMO1_5  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DQWTVQPIQL PEKESWTVND
 96BWMO3_2  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL AEKDSWTVND
98BWMC12_2  IHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTIQPIQL PEKDSWTVND
98BWMC13_4  RHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKESWTVND
98BWMC14_a  DHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PTKDSWTVND
98BWMO14_1  GHLLQWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
98BWMO18_d  EHLLKWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
98BWMO36_a  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
98BWMO37_d  QHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
99BW3932_1  DHLLGWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKESWTVND
99BW4642_4  GHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
99BW4745_8  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
99BW4754_7  EHLLRWGLTT PDKKHQKEPP FLWMGYEPHP DKWTVQPILL PEKESWTVND
99BWMC16_8  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIML PTKDSWTVND
A2_CD_97CD  AHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIKL PEKDSWTVND
A2_CY_94CY  AHLLKWGFYT PDKKHQKEPP FLWMGYELHP DKWTVQPIKL PEKDSWTVND
A2D___97KR  NHLLKWGFYT PDKKHQKEPP FLWMEYELHP DKWTVQPIKL PEKDSWTVND
A2G_CD_97C  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKESWTVND
A_BY_97BL0  AHLLSWGFTT PDKKHQKEPP FLWXXYEXHP DKWTVQPIML PDKDSWTVND
A_KE_Q23_A  AHLLSWGLIT PDKKHQKEPP FLWMGYELHP DKWTVQPIEL PEKDSWTVND
A_SE_SE659  SHLLSWGFTT PDQKHQKEPP FLWMGYELHP DKWTVQPIVL PEKESWTVND
A_SE_SE725  AHLLSWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIEL PEKESWTVND
```

```
A_SE_SE753   EHLLSWGFTT PDKKHQKEPP FLWMGYEVHP DKWTVQPIVL PEKESWTVND
A_SE_SE853   AHLLSWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPINL PEKESWTVND
A_SE_SE889   AHLLSWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PNKENWTVND
A_SE_UGSE8   AHLLSWGFFT PDQKHQKEPP FLWMGYELHP DKWTVQPIKL PEKESWTVND
A_UG_92UG0   EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIEL PEKESWTVND
A_UG_U455_   AHLLSWGFIT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKESWTVND
AC_IN_2130   EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
AC_RW_92RW   EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
AC_SE_SE94   AHLLSWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
ACD_SE_SE8   AHLLSWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
ACG_BE_VI1   EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQSIKL PEKESWTVND
AD_SE_SE69   GHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
AD_SE_SE71   EHLLKWGFYT PDQKHQKEPP FLWMGYELHP DKWTVQPITL PEKESWTVND
ADHK_NO_97   EHLLRWGLTT PDKKHQKEPP FLWMGYELHP DRWTVQPIKL PEKESWTVND
ADK_CD_MAL   EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKESWTVND
AG_BE_VI11   EHLLRWGLTT PDQKHQKEPP FLWMGYELHP DKWTVQPIQL PDKESWTVND
AG_NG_92NG   NHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
AGHU_GA_VI   AHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQTVKL PEKDSWTVND
AGU_CD_Z32   EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKESWTVND
AJ_BW_BW21   EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKEAWTVND
B_AU_VH_AF   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_CN_RL42_   QHLLRWGFTT PDKKHQKEPP FLWMGYELHL DKWTVQPIML PEKDSWTVND
B_DE_D31_U   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPITL PEKDSWTVND
B_DE_HAN_U   QHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_FR_HXB2_   QHLLRWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_GA_OYI__   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIML PEKDSWTVND
B_GB_CAM1_   QHLLRWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIML PEKDSWTVND
B_GB_GB8_A   QHLLRWGFST PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_GB_MANC_   QYLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPITL PEKDSWTVND
B_KR_WK_AF   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_NL_3202A   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_TW_TWCYS   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_US_BC_L0   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_US_DH123   QHLLRWGLFT PDQKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_US_JRCSF   QHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_US_MNCG_   RHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_US_P896_   QHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_US_RF_M1   EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
B_US_SF2_K   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIML PEKDSWTVND
B_US_WEAU1   QHLLRWGFTT PDKKHQKDPP FLWMGYELHP DKWTVQPIKL PEKESWTVND
B_US_WR27_   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIEL PEKDSWTVND
B_US_YU2_M   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
BF1_BR_93B   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
C_BR_92BR0   EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
C_BW_96BW0   EHLLKWGLTT PYKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVHD
C_BW_96BW1   KHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQSIKL PEKESWTVND
C_BW_96BW1   EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKESWTVND
C_BW_96BW1   KHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIEL PEKESWTVND
C_ET_ETH22   EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
C_IN_93IN1   EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
C_IN_93IN9   GHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
C_IN_93IN9   QHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
C_IN_94IN1   EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
C_IN_95IN2   KHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
CRF01_AE_C   DHLLKWGFTT PDKKHQKEPP FLWMGYELHP DRWTVQPIEL PEKDSWTVND
CRF01_AE_C   AHLLSWGFTT PDKKHQKEPP FLWMGYELHP DRWTVQPIVL PEKDSWTVND
CRF01_AE_C   AHLLSWGFTT PDKKHQKESP FLWMGYELHP DRWTVQPIQL PDKESWTVDD
CRF01_AE_T   AHLLSWGFTT PDKKHQKEPP FLWMGYELHP DRWTVQPIEL PEKDSWTVND
CRF01_AE_T   AHLLSWGFTT PDKKHQKEPP FLWMGYELHP DRWTVQPIEL PEKDSWTVND
CRF01_AE_T   AHLLSWGFTT PDKKHQKEPP FLWMGYELHP DRWTVQPIEL PEKDSWTVND
CRF01_AE_T   AHLLSWGFTT PDQKHQKEPP FLWMGYELHP DRWTVQPIEL PEKDSWTVND
```

```
CRF01_AE_T  AHLLSWGFTT PDKKHQKEPP FLWMGYELHP DRWTVQPIEL PEKDSWTVND
CRF01_AE_T  AHLLSWGFTT PGKKHQKEPP FLWMGYELHP DRWTVQPIEL PEKDSWTVND
CRF02_AG_F  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
CRF02_AG_F  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
CRF02_AG_G  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIML PEKDSWTVND
CRF02_AG_N  GHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPVEL PEKDSWTVND
CRF02_AG_S  EHLLRWGFTT PDEKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
CRF02_AG_S  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
CRF03_AB_R  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
CRF03_AB_R  DHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIVL PEKDSWTVND
CRF04_cpx_  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQP AEKDSWTVND
CRF04_cpx_  EHLLRWGFYT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
CRF04_cpx_  EHLLRWGFFT PDQKHQKEPP FLWMGYELHP DKWTVQPIQL VEKESWTVND
CRF05_DF_B  EHLLAWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
CRF05_DF_B  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PNKDSWTVND
CRF06_cpx_  EHLLKWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKESWTVND
CRF06_cpx_  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PNKESWSIND
CRF06_cpx_  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
CRF06_cpx_  AHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
CRF11_cpx_  KHLLRWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKECWTVND
CRF11_cpx_  KHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKESWTVND
D_CD_84ZR0  EHLLRWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQSITL PEKESWTVND
D_CD_ELI_K  EHLLRWGFTR PDKKHQKEPP FLWMGYELHP DKWTVQSIKL PEKESWTVND
D_CD_NDK_M  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPINL PEKESWTVND
D_UG_94UG1  GHLLKWGFTT PDKKYQKEPP FLWMGYELHP DKWTVQPIHL PEKESWTVND
F1_BE_VI85  EHLLRWGFTT PDKKHQKEPP FLWMGHELHP DKWTVQPIQL PNKDSWTVND
F1_BR_93BR  EHLLKWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
F1_FI_FIN9  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVDD
F1_FR_MP41  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKDSWTVND
F2_CM_MP25  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQAIQL PDKSSWTVND
F2KU_BE_VI  EHLLRWGFFT PDEKHQKEPP FLWMGYELHP DKWTVQPIQL PDKEDWTVND
G_BE_DRCBL  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PNKENWTVND
G_NG_92NG0  EHLLKWGLTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKEDWTVND
G_SE_SE616  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKESWTVND
H_BE_VI991  AHLLRWGFTT PDQKHQKEPP FLWMGYELHP DKWTVQPVKL PEKDSWTVND
H_BE_VI997  AHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPVKL PEKDSWTVND
H_CF_90CF0  AHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQTVKL PEKDSWTVND
J_SE_SE702  EHLLKWGFYT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKEDWTVND
J_SE_SE788  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PEKEDWTVND
K_CD_EQTB1  EHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
K_CM_MP535  EHLLKWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKDSWTVND
N_CM_YBF30  DHLLKWGFTT PDKKHQKEPP FLWMGYELHP. DKWTVQPIKL PEKDVWTVND
O_CM_ANT70  EHLYQWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQSIQL PNKDVWTVND
O_CM_MVP51  EHLYQWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PDKEVWTVND
O_SN_99SE_  EHLYQWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PNKEEWTVND
O_SN_99SE_  EHLYQWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PNKEEWTVND
U_CD___83C  AHLLRWGFTT PDKKHQKEPP FLWMGYELHP DKWTVQPIQL PNKDDWTVND

            451                                                 500
00BW0762_1  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW0768_2  IQKLVGKLNW ASQIYPGIKV RQLCRLLRGA KVLTDIVPLT EEAELELAEN
00BW0874_2  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGT KALTDIVPLT EEAELELAEN
00BW1471_2  IQKLVGKLNW ASQIYAGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW1616_2  IQKLVGKLNW ASQIYPGVKV KQLCKLLRGT KALTDIVPLT EEAELELAEN
00BW1686_8  IQKLVGKLNW ASQIYSGIKV KQLCKLLRGT KALTDIVPLT EEAELELAEN
00BW1759_3  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW1773_2  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW1783_5  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW1795_6  IQKLVGKLNW ASQIYPGIKV TQLCKLLRGA KALTDIVPLT DEAELELAEN
00BW1811_3  IQKLVGKLNW ASQIYSGIKV RQLCKLLRGT KALTDIVPLT DEAELELAEN
00BW1859_5  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
```

```
00BW1880_2   IQKLVGKLNW ASQIYKGIKV KQLCKLLRGA KALTDIVPLP EEAELELAEN
00BW1921_1   IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTDIVPLT EEAELELAEN
00BW2036_1   IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT EEAEVELAEN
00BW2063_6   IQKLVGKLNW ASQIYPGIKV TQLCKLLRGA KALTDIVPLT EAAELELAEN
00BW2087_2   IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTDIVPLT EEAELELAEN
00BW2127_2   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW2128_3   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGT KALTDIIPLT EEAELELAEN
00BW2276_7   IQKLVGKLNW VSQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW3819_3   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW3842_8   IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTDIIPLT EEAELELAEN
00BW3871_3   LQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTDIVPLT EEAELELAEN
00BW3876_9   IQKLVGKLNW ASQIYPGIKV RHLCRLLRGA KALTEIVPLT EEAELELAEN
00BW3886_8   IQKLVGKLNW ASQIYTGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW3891_6   IQKLVGKLNW ASQIYAGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
00BW3970_2   IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTDIVPLT EEAELELAEN
00BW5031_1   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
 96BW01B21   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
  96BW0407   IQKLVGKLNW ASQIYPGIRV RHLCKLLRGA KALTDIVPLT EEAELELAEN
  96BW0502   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDVVPLT EEAELELAEN
 96BW06_J4   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGT KALTDVVPLN EEAELELAEN
 96BW11_06   IQKLVGKLNW ASQIYPGVKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
  96BW1210   IQKLVGKLNW ASQIYPGIKV RQLCKILRGV KALTDIVTLT EEAELELAEN
 96BW15B03   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAED
 96BW16_26   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIIPLT EEAELELAEN
 96BW17A09   IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
 96BWMO1_5   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT DEAELELAEN
 96BWMO3_2   VQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
98BWMC12_2   IQRLVGKLNW ASQIYSGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
98BWMC13_4   IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
98BWMC14_a   IQKLVGKLNW ASQNYPGIKV RQLCKLLRGT KALTDIVPLT EEAELELAEN
98BWMO14_1   IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
98BWMO18_d   IQKLVGKLNW ASQIYPGIKV RQLCKLIRGT KKLTDIVPLT EEAELELAEN
98BWMO36_a   IQKLVGKLNW ASQIYPGIKV KNLCKLLRGA KALTDIVPLT EEAELELAEN
98BWMO37_d   IQELVGKLNW ASQIYSGIKV KQLCKLLRGT KALTDIVPLT EEAELELAEN
99BW3932_1   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
99BW4642_4   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
99BW4745_8   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGT KALTDIVPLT EEAELELAEN
99BW4754_7   IQKLVGKLNW ASQIYPGIQV KQLCKLLRGA KALTDIVPLT EEAELELAEN
99BWMC16_8   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
A2_CD_97CD   IQKLVGKLNW ASQIYAGIKV RQLCKLLRGT KALTDIVPLT REAELELEEN
A2_CY_94CY   IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVTLT KEAELELEEN
A2D___97KR   IQKLVGKLNW ASQIYAEIKV KQLCKLLRGA KALTDIVPLT KEAELELEEN
A2G_CD_97C   IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTEIVSLT AEAELELAEN
A_BY_97BL0   IQKLXGKLNW ASQIYPEIKV RQLCKLLXGA KALTDIVTLT EEAELELAEN
A_KE_Q23_A   IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDVVTLT EEAELELAEN
A_SE_SE659   IQKLVGKLNW ASQIYPGIKV KQLCRLLRGA KALTDIVELT EEAELELAEN
A_SE_SE725   IQKLVGKLNW ASQIYAGIKV KQLCRLLRGA KALTDIVTLT EEAELELAEN
A_SE_SE753   IQKLVGKLNW ASQIYAGIKV RQLCKLLRGA KALTDIVTLT EEAELELAEN
A_SE_SE853   IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVTLT EEAELELAEN
A_SE_SE889   IQKLVGKLNW ASQIYVGIKV KQLCKLLRGA KALTDIVTLT EEAELELAEN
A_SE_UGSE8   IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVTLT EEAELELAEN
A_UG_92UG0   IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTDIVTLT EEAELELAEN
A_UG_U455_   IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVTLT EEAELELAEN
AC_IN_2130   IQKLVGKLNW ASQIYPGIKV RQLCRLLRGA KALTDIVPLT EEAELELAEN
AC_RW_92RW   IQKLVGKLNW ASQIYPGVKV RQLCKLLRGT KALTDIVPLT EEAELELAEN
AC_SE_SE94   IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KSLTDVVTLT EEAELELAEN
ACD_SE_SE8   IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDVVTLT EEAELELAEN
ACG_BE_VI1   IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
AD_SE_SE69   IQKLVGKLNW ASQIYPGIKV RQLCKCIRGA KALTEVVPLT EEAELELAEN
AD_SE_SE71   IQKLVGKLNW ASQIYAGIKV KQLCKCLRGA KALTEIVPLT EEAELELAEN
ADHK_NO_97   IQKLVGKLNW ASQIYPGIKV RQLCRLLRGT KALTDIVPLT AEAELELAEN
```

```
ADK_CD_MAL  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT AEAELELAEN
AG_BE_VI11  IQKLVGKLNW ASQIYAGIRV KQLCKLLRGA KALTDIVTLT EEAELELAEN
AG_NG_92NG  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVTLT EEAELELAEN
AGHU_GA_VI  IQKLVGKLNW ASQIYSGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
AGU_CD_Z32  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTDIVPLT AEAELELAEN
AJ_BW_BW21  IQKLVGKLNW ASQIYPGIQV RHLCRLLRGA KALTDIVPLT AEAELELAEN
B_AU_VH_AF  IQKLVGKLNW ASQIYSGIKV RQLCKLLRGT KALTEVIPLT EEAELELAEN
B_CN_RL42_  IQKLVGKLNW ASQIYAGIKV KELCKLLRGT KALTEVIPLT EEAELELAEN
B_DE_D31_U  IQKLVGKLNW ASQIYAGIKV RQLCKLLRGA KALTEVIPLT KEAELELAEN
B_DE_HAN_U  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGP KALTEVIPLT KEAELELAEN
B_FR_HXB2_  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGT KALTEVIPLT EEAELELAEN
B_GA_OYI__  IQKLVGKLNW ASQIYAGIKV KNLCKLLRGT KALTEVIPLT EEAELELAEN
B_GB_CAM1_  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTEVVPLT EEAELELAEN
B_GB_GB8_A  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTEVITLT EEAELELAEN
B_GB_MANC_  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTEVIPLT KEAELELAEN
B_KR_WK_AF  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTEVVPLT EEAELELAEN
B_NL_3202A  IQKLVGKLNW ASQIYAGIKV RHLCKLLRGT KALTEVIPLT EEAELELAEN
B_TW_TWCYS  IQKLVGKLNW ASQIYAGIKV RQLCKLLRGT KALTEVVPLT KEAELELAEN
B_US_BC_L0  IQKLVGKLNW ASQIYSGIKV KQLCKLLRGT KALTEVVTLT EEAELELAEN
B_US_DH123  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTEVIPLT EEAELELAEN
B_US_JRCSF  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTEVIPLT KEAELELAEN
B_US_MNCG_  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTEVIPLT EEAELELAEN
B_US_P896_  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTEVVPLT EEAELELAEN
B_US_RF_M1  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTEVVQLT KEAELELAEN
B_US_SF2_K  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTEVIPLT EEAELELAEN
B_US_WEAU1  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTEIIPIT EEAELELAEN
B_US_WR27_  IQKLVGKLNW XSQIYAGIKV XQLCKLLRGT KALTEVVPLT EEAELELAGN
B_US_YU2_M  IQKLVGKLNW ASQIYAGIKV RQLCKLLRGT KALTEVIPLT EEAELELAEN
BF1_BR_93B  IQKLVGKLNW ASQIYAGIKV RQLCKLLRGT KALTEVVPLT AEAELELAEN
C_BR_92BR0  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
C_BW_96BW0  IQKLVGKLNW ASQIYPGIRV KHLCKLLRGA KALTDIVPLT EEAELELAEN
C_BW_96BW1  IQKLVGKLNW ASQIYPGVKV RQLCKLLRGA KALTDIVPPT EEAELELAEN
C_BW_96BW1  IQKLVGKLNW ASQIYPGIKV RQLCKILRGV KALTDIVTLT EEAELELAEN
C_ET_ETH22  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAED
C_IN_93IN1  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVTLT EEAELELAEN
C_IN_93IN9  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
C_IN_93IN9  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
C_IN_94IN1  IQKLVGKLNW ASQIYPGIKV KQLCRLLRGA KVLTDIVPLT EEAELELAEN
C_IN_95IN2  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT EEAELELAEN
CRF01_AE_C  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGT KALTDIVPLT EEAELELAEN
CRF01_AE_C  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTEVVPLT EEAELELAEN
CRF01_AE_C  IQKLVGKLNW ASQIYAGIKV RQLCKLLRGA KALTDIVTLT EEAELELAEN
CRF01_AE_T  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
CRF01_AE_T  IQKLVGKLNW ASQIYAGIKV KQLCKPLRGT KALTDIVPLT EEAELELAEN
CRF01_AE_T  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
CRF01_AE_T  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTDIVPLT EEAELELEEN
CRF01_AE_T  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
CRF01_AE_T  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
CRF02_AG_F  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDTVPLT EEAELELAEN
CRF02_AG_F  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVPLT EEAELELAEN
CRF02_AG_G  IQKLVGKLNW ASQIFAGIKV KQLCRLLRGA KALTDIVPLT EEAELELAEN
CRF02_AG_N  IQKLVGKLNW ASQIYAGIKV KELCKLLRGA KALTDIVTLT EEAELELAEN
CRF02_AG_S  IQKLVGKLNW ASQIYAGIKI KQLCRLLRGA KALTDIVALT EEAELELAEN
CRF02_AG_S  IQKLVGKVNW QVRIYAGIKV KQLCKLLRGA KTLTDIVTLT EEAELELAEN
CRF03_AB_R  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTDIVPLT EEAELELAEN
CRF03_AB_R  IQKLVGKLNW ASQIYAGIKV RQLCKLLRGA KALTEVIPLT AEAELELAEN
CRF04_cpx_  IQKLVGKLNW ASQIYAGIKV RQLCKLLRGA KALTEIIPLT AEAELELAEN
CRF04_cpx_  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT TEAELELAEN
CRF04_cpx_  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTDIVPLT TEAELELAEN
CRF05_DF_B  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVPLT TEAELELAEN
            IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTEVIPLT EEAELELAEN
```

```
CRF05_DF_B  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTEVVPLT EEAELELAEN
CRF06_cpx_  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDIVPLT AEAELELAEN
CRF06_cpx_  IQKLVGKLNW ASQIYSGIKV RQLCKLLRGA KALTDIVPLT AEAELELAEN
CRF06_cpx_  IQKLVGKLNW ASQIYPGIKV KHLCKLLRGA KALTDIVPLT AEAELELAEN
CRF06_cpx_  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT AEAELELAEN
CRF11_cpx_  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTDIVPLT AEAELELAEN
CRF11_cpx_  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGA KALTDIVPLT AEAELELAEN
D_CD_84ZR0  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGT KALTEVIPLT EEAELELAEN
D_CD_ELI_K  IQNLVERLNW ASQIYPGIKV RQLCKLLRGT KALTEVIPLT EEAELELAEN
D_CD_NDK_M  IQKLVGKLNW ASQIYAGIKV KQLCKLLRGT KALTEVVPLT EEAELELAEN
D_UG_94UG1  IQKLVGKLNW ASQIYPGIKV RQLCKCLRGA KALTEVIPLT AEAELELAEN
F1_BE_VI85  IQKLVGKLNW ASQIYPGIKV RPLCKLLRGA KALTDIVPLT AEAELELAKN
F1_BR_93BR  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGA KALTDIVPLT TEAELELAEN
F1_FI_FIN9  IQKLVGKLNW ASXIYPGIKV RQLCKLLRGA KALTDMVPLT AEANLELAEN
F1_FR_MP41  IQKLVGKLNW ASQIYPGIKI KQLCKLLRGA KALTDIVPLT EEAELELAEN
F2_CM_MP25  IQKLVGKLNW ASQIYPGIRV KHLCKLLRGT KALTDVVPLT AEAELELAEN
F2KU_BE_VI  IQKLVGKLNR ASQIYPGIKV KQLCKLIRGA KTLTDIVPLT AEAELELAEN
G_BE_DRCBL  IQKLVGKLNW ASQIYPGIKV KQLCKLIRGA KALTDIVSMT AEAEMELAEN
G_NG_92NG0  IQKLVGKLNW ASQIYPGIKV KHLCRLLRGA KALTDIVPLT AEAEMELAEN
G_SE_SE616  IQKLVGKLNW ASQIYPGIKV THLCKLLRGA KALTDIVSLT AEAEMELAEN
H_BE_VI991  IQKLVGKLNW ASQIYPGIKV KQLCXLLRGA KALTEIVPLT KEAELELAEN
H_BE_VI997  IQKLVGKLNW ASQIYPGIKV RQLCKLLRGA KALTDVVPLT KEAELELAEN
H_CF_90CF0  IQKLVGKLNW ASQIYPNIKV KQLCKLLRGA KALTDIIPLT KEAELELAEN
J_SE_SE702  IQKLVGKLNW ASQIYPGIKI KELCKLIRGA KALTDIVPLT REAELELAEN
J_SE_SE788  IQKLVGKLNW ASQIYPGIKV KQLCKLLKGA KALTDIVPLT REAELELAEN
K_CD_EQTB1  IQKLVGKLNW ASQIFPGIKV KQLCKLLRGV KALTDIVPLT AEAELELAEN
K_CM_MP535  IQKLVGKLNW ASQIYPGIKV KQLCKLLRGV KALTDIVPLT AEAELELAEN
N_CM_YBF30  IQKLVGKLNW ASQIYPGIRV KQLCKLIRGT KALTEVVNFT EEAELELAEN
O_CM_ANT70  IQKLIGKLNW ASQIYQGIRV RELCKLIRGT KSLTEVVPLS REAELELEEN
O_CM_MVP51  IQKLVGKLNW ASQIYQGIRV KELCKLIRGT KSLTEVVPLS KEAELELEEN
O_SN_99SE_  IQKLVGKLNW ASQIYQGIRV KELCKLIRGA KSLTEIVPLS KEAELELEEN
O_SN_99SE_  IQKLVGKLNW ASQIYQGIRV KELCKLIRGT KSLTEVVPLS KEAELELEEN
U_CD___83C  IQKLVGKLNW ASQIYPGIQV RQLCKLLRGT KALTDIVPMT KEAELELAEN

            501                                                 550
00BW0762_1  R.EILKEPVH GVYYDPSKDL IADIQKQGND QWTYQIYQEP FKNLKTGKYA
00BW0768_2  R.EILREPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
00BW0874_2  R.EILREPVH GVYYDPSKDL VAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
00BW1471_2  R.KILKEPVH GVYYDPSKDL IAEIQKQGHE QWTYQIYQEP FKNLKTGKYA
00BW1616_2  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW1686_8  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW1759_3  R.EILKEPVH GVYYDPSKEL IAELQKQGND QWTYQIYQEP FKNLKTGKYA
00BW1773_2  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW1783_5  R.EILKEPVH GVYYDPSKDL KAEIQKQGND QWTYQIYQEP FKNLKTGKYA
00BW1795_6  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW1811_3  R.EILKEPVH GAYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW1859_5  R.EILKEPVH EVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW1880_2  R.EILKEPVH GVYYDPSKDL IAEVQKQGDD QWTYQIYQEP LKNLKTGKYA
00BW1921_1  R.EILKEPVH GVYYDPSKEL IAEIQKQGQD QWSYQIYQEP FKNLKTGKYA
00BW2036_1  R.EILRGPVH GVYYDPSKDL VAEIQKQGHD QWTYQIFQEP FKNLKTGKYA
00BW2063_6  R.EILREPVH GVYYDPSKDL VAEIQKQGND QWTYQIYQEP FKNLKTGKYA
00BW2087_2  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW2127_2  R.EILKEPVH GVYYDPSKDL IAEIQKQGND QWTYQIYQEP FKNLKTGKYA
00BW2128_3  R.EILKEPVH GVYYDSSKEL IAEIQKQGHG QWTYQIYQEP FKNLKTGKYA
00BW2276_7  R.EILKEPVH GVYYDPSEDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW3819_3  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW3842_8  R.EILREPVH GVYYDPSKDL VAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
00BW3871_3  R.EILKEPVH GVYYDPSKDL IAEIQKQGND QWTYQIYQEP FKNLKTGKYA
00BW3876_9  RXEILKEPVH GVYYDPSKDL IAEIQKQGYD QWTYQIYQEP YKNLKTGKYA
00BW3886_8  R.EILKGPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
00BW3891_6  R.EILKEPVH GVYYDPSKDL IAEIQKQGNG .QTYQIYQEP FKNLKTGKYA
```

```
00BW3970_2  R.EILREPVH GVYYDPSKDL IAEIQKQGND QWTYQIYQEP FKNLKTGKYA
00BW5031_1  R.EILREPVH GVYYDPSKDL IAEIQKQGDD QWTYQIYQEP FKNLKTGKYA
 96BW01B21  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
  96BW0407  R.EILKEPVH GVYYDPSEDL IAEVQKQGHY QWTYQIYQEP FKNLKTGKYA
  96BW0502  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
 96BW06_J4  R.EILKEPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
 96BW11_06  R.EILKEPVH GVYYDPSKDL IAEIQKQGGD QWTYQIYQEP FKNLKTGKYA
  96BW1210  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
 96BW15B03  R.EILREPVH GVYYDPSKDL VAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
 96BW16_26  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
 96BW17A09  R.EILKEPVH GAYYDPSKDL IAEIQKQGNG QWTYQIYQEP FKNLKTGKYA
 96BWMO1_5  R.EILREPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
 96BWMO3_2  R.EILKEPVH GVYYDPSKEL IAEIQKQGDD QWTYQIYQEP HKNLKTGKYA
98BWMC12_2  R.EILKEPVH GVYYDPSKDL IAEIQKQGYD QRTYQIYQEP FKNLKTGKYA
98BWMC13_4  R.EILKEPVH GVYYDPSKDL VAEIQKQGLD QWTYQIYQEP FKNLKTGKYA
98BWMC14_a  R.EILKEPVH GVYYDPSKDL IAEIQKQGND QWTYQIYQEP FKNLKTGKYA
98BWMO14_1  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
98BWMO18_d  R.EILKEPVH GVYYDPSKDL IAEIQKQGDD QWTYQIYQEP FKNLKTGKYA
98BWMO36_a  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
98BWMO37_d  R.EILKEPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
99BW3932_1  R.ELLKEPVH GVYYDPSKDL ITEIQKQGHD QWTYQIYQEP FKNLKTGKYA
99BW4642_4  R.EILKEPVH GVYYDPSKDL IAEVQKQGCE QWTYQIYQEP FKNLKTGKYA
99BW4745_8  R.EILKEPVH GVYYDPSKDL IAEVQKQGND QWTYQIYQEP FKNLKTGKYA
99BW4754_7  R.EILKEPVH GVYYDPSKDL IAEIQKQGNG QWTYQIYQEP FKNLKTGKYA
99BWMC16_8  R.EILKEPVH GVYYDPSKEL IAEIQKQGND QWTYQIYQEP FKNLKTGKYA
A2_CD_97CD  R.EILKDPVH GVYYDPSKDL IAEIQKQGQG QWSYQIYQEP FKNLKTGKYA
A2_CY_94CY  R.EILKTPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
A2D___97KR  R.EILKDPVH GVYYDPSKDL IAEVQKQGPD QWTYQIYQEP FKNLKTGKYA
A2G_CD_97C  R.EILKEPVH GAYYEPSKEL IAEVQKQGLD QWTYQIYQEP YKNLKTGKYA
A_BY_97BL0  R.EILKXPVH XVYYDPSKDL VAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
A_KE_Q23_A  R.EILKEPVH GVYYDPSKDL IAEIQKQGQD QWTYQIFQEP FKNLKTGKYA
A_SE_SE659  R.EILKDPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
A_SE_SE725  R.EILKDPVH GVYYDPSKDL IAEIQKQGLD QWTYQIYQEP FKNLKTGKYA
A_SE_SE753  R.EILKDPVH GAYYDPSKDL IVEIQKQGQD QWTYQIYQET FKNLKTGKYA
A_SE_SE853  R.EILKAPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
A_SE_SE889  R.EILKDPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
A_SE_UGSE8  R.EILKDPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
A_UG_92UG0  R.EILKDPVH GAYYDPSKDL IAEIQKQGQD QWIYQIYQEP FKNLKTGKYA
A_UG_U455_  R.EILKDPVH GVYYDPSKDL VAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
AC_IN_2130  R.EILKEPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
AC_RW_92RW  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
AC_SE_SE94  R.EILRDPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
ACD_SE_SE8  R.EILKDPVH GVYYDPSKDL IAEIQKQGQE QWTYQIYQEP FKNLKTGKYA
ACG_BE_VI1  R.EILKEPVH GVYYDPAKDL VAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
AD_SE_SE69  R.EILKEPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEQ YKNLKTGKYA
AD_SE_SE71  R.EILKEPVH GVYYDPSKDL IAEIQKQGQD QWSYQIYQEQ YKNLKTGKYA
ADHK_NO_97  R.EILKEPVH GVYYDPSKDL IAEIQKQGRG QWTYQIYQEP YKNLKTGKYA
ADK_CD_MAL  R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEQ YKNLKTGKYA
AG_BE_VI11  R.EILKEPVH GVYYDPGKEL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
AG_NG_92NG  R.EILKEPVH GVYYDPSKEL IAELQKQGCD QWTYQIYQEP YKNLKTGKYA
AGHU_GA_VI  R.EILKEPVH GVYYDPSKDL VAEVQKQGPD QWTYQIYQEP FKILKTGKYA
AGU_CD_Z32  R.EILKEPVH GVYYDSSKEL IAEVQKQGLN QWTYQIYQEP FKNLKTGKYA
AJ_BW_BW21  R.EILKEPVH GVYYDSAKEL IAEVQKQGLD QWTYQIYQEP FKNLKTGKYA
B_AU_VH_AF  R.EILKEPVH GVYYDPSKDL IAEVQKQEQG QWTYQIYQEP FKNLKTGKYA
B_CN_RL42_  R.EILKESVH GVYYDPSKDL IAEIQKQGLG QWTYQIYQEP YKNLKTGKYA
B_DE_D31_U  R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
B_DE_HAN_U  R.EILKEPVH GVYCDPSKDL VAEIQKQGEG QWTYQIYQEP FKNLKTGKYA
B_FR_HXB2_  R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
B_GA_OYI__  R.EILKEPVH GVYYDPSKDL VAELQKQGQG QWTYQIYQEP FKNLKTGKYA
B_GB_CAM1_  R.EILKEPVH GVYYDPSKDL IAELQKQGQG QWTYQIYQEP FKNLKTGKYA
B_GB_GB8_A  R.EILKEPVH GVYYDPSKDL VAEIQKQGLG QWTYQIYQEP FKNLKTGKYA
```

```
B_GB_MANC_  R.EILKVPVH GVYYDPSKDL TAEIQKQGQG QWTYQIYQEP FRSLKTGKYA
B_KR_WK_AF  R.EILKEPVH GVYYDPAKDL IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
B_NL_3202A  R.EILKEPVH GVYYDPSKEL IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
B_TW_TWCYS  R.EILKQPVH GAYYDPSKDL VAEIQKQGQG QWTYQIYQET FKNLKTGKYA
B_US_BC_L0  R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEP FKNLKTAKYA
B_US_DH123  R.EILKEPVH GVYYDPSKDI IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
B_US_JRCSF  R.EILKEPVH GVYYDPSKDL IVEIQKQGQG QWTYQIFQEP FKNLKTGKYA
B_US_MNCG_  R.EILKEPVH GVYYDPSKDL IAEVQKQGQG QWTYQIYQEP FKNLKTGKYA
B_US_P896_  R.EILKEPVH GVYYDPTKDL IAELQKQGQG QWTYQIYQEP YKNLKTGKYA
B_US_RF_M1  R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
B_US_SF2_K  R.EILKEPVH EVYYDPSKDL VAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
B_US_WEAU1  R.EILKEPVH GVYYDPSKDL IAELQKQGQG QWTYQIYQEP FKNLKTGKYA
B_US_WR27_  R.EILKEPXH GVYYDPSXDL VAELQKQGXG QWTYQIYQEP FIXLXTGKYA
B_US_YU2_M  R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
BF1_BR_93B  R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEP YKNLKTGKYA
C_BR_92BR0  R.EILKEPVH GVYYDPSKDL IAEIQKQGQN QWTYQIYQEP FKNLKTGKYA
C_BW_96BW0  R.EILKEPVH GVYYDPSKDL IAEVQKQGHD QWTYQIYQEP FKNLKTGKYA
C_BW_96BW1  R.EILKEPVH GVYYDPSKDL IAEIQKQGGD QWTYQIYQEP FKNLKTGKYA
C_BW_96BW1  R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
C_BW_96BW1  R.EILREPVH GVYYDPSKDL VAEIQKQGHD QWTYQIYQEP FKNLKTGKYA
C_ET_ETH22  R.EILKEPVH GVFYDPSKDL IAEIQKQGND QWTFQFYQEP FKNLKTGKFA
C_IN_93IN1  R.EILKEPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
C_IN_93IN9  R.EILKEPVH GVYYDPSKDL IAEIQKQGQD QWTYQVYQEP FKNLKTGKYA
C_IN_93IN9  R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
C_IN_94IN1  R.EILKEPIH GVYYDPSKDL IAEIQKQGQD QWTYQVYQEP FKNLKTGKYA
C_IN_95IN2  R.EILKEPVH GVYYDPSKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF01_AE_C  R.EILKNPVH GVYYDPSKEL VAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF01_AE_C  R.EILKTPVH GVYYDPSKDL VAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF01_AE_C  R.EILKNPVH GVYYDPSKDL VAEIQKQGHN QWTYQIYQEP FKNLKTGKYA
CRF01_AE_T  R.EILKTPVH GVYYDPSKDL VAEVQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF01_AE_T  R.EILRTPVH GVYYDPSKDL VAEVQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF01_AE_T  R.EILRIPVH GVYYDPSKDL VAEVQKQGQD QWTYQIYQEP FKNLKTGKYS
CRF01_AE_T  R.EILKTPVH GVYYDPSKDL AAEVQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF01_AE_T  R.EILKTPVH GVYYDPSKDL VAEVQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF01_AE_T  R.EILKTPVH GVYYDPSKDL VAEVQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF02_AG_F  R.EILKEPVH GVYYDPAKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF02_AG_F  R.EILKEPVH GVYYDPAKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF02_AG_G  R.EILKEPVH GVYYDPTKDL IAEIQKQGQD QWTFQIYQEP FKNLKTGKYQ
CRF02_AG_N  R.EILKEPVH GVYYDPTKDL VAELQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF02_AG_S  R.EILKEPVH GVYYDPTKDL VAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF02_AG_S  R.EILKEPVH GVYYDPTKDL IAEIQKQGQD QWTYQIYQEP FKNLKTGKYA
CRF03_AB_R  R.EILKEPVH GVYYDPSKDL VAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
CRF03_AB_R  R.EILKVPVH GVYYDPSKDL VAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
CRF04_cpx_  R.EILKEPVH GAYYDPSKDL IAEIQKQGQG QWTYQIYQEP HKNLKTGKYA
CRF04_cpx_  R.EILKEPVH GAYYDPSKDL IAKIQKQGQG QWTYQIYQEP YKNLKTGKYA
CRF04_cpx_  R.EILKEPVH GAYYDPSKDL IAEIQKQGLG QXTYQIYQEP YKNLKTGKYA
CRF05_DF_B  R.EILKEPVH GVYYDPSKDL IAEVQKQGLG QWTYQIYQEP FKNLKTGKYA
CRF05_DF_B  R.EILKEPVH GVYYDPAKDL IAEIQKQGQE QWTYQIYQEP FKNLKTGKYA
CRF06_cpx_  R.EILKEPVH GVYYDPSKDL IAEVQKQGYG QWTYQIYQEP HKNLKTGKYA
CRF06_cpx_  R.EILKEPVH GAYYDPSKDL IAEIQKQGQG QWTYQIYQEP HKNLKTGKYA
CRF06_cpx_  R.EILKEPVH GVYYDPLKDL IAELQKQGQG QWTYQIYQEP HKNLKTGKYA
CRF06_cpx_  R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQDT HKNLKTGKYA
CRF11_cpx_  R.EILKEPVH GVYYDPSKDL IAEVQKQGLD QWTYQIYQEP FKNLKTGKYA
CRF11_cpx_  R.EILKEPVH GVYYDPSKDL IAEVQKQGLD QWTYQIYQEP FKNLKTGKYA
D_CD_84ZR0  R.EILKEPMH GVYYDPSKDL IAELQKQGQG QWTYQIYQEP FKNLKTGKYA
D_CD_ELI_K  R.EILKEPVH GVYYDPSKDL IAEIQKQGHG QWTYQIYQEP FKNLKTGKYA
D_CD_NDK_M  R.EILKEPVH GVYYDPSKDL IAELQKQGDG QWTYQIYQEP FKNLKTGKYA
D_UG_94UG1  R.EILKEPVH GAYYDPSKDL IAEIQKQGQD QWTYQIYQEQ YKNLKTGKYA
F1_BE_VI85  R.EILREPVH GVYYDPSKDL IAEIQKQGDG QWTYQIYQNP FKNLKTGKYA
F1_BR_93BR  R.EILKEPVH GAYYDPSKDL IAEIQKQGQG QWTYQIYQEP FKNLKTGKYA
F1_FI_FIN9  R.EILKEPVH GVYYDPSKDL IPKLQKQGQG QWTYQIYREP FKNLKTGKYA
```

```
F1_FR_MP41   R.EILKEPVH GVYYDPSKDL IAELQKQGQG QWTYQIYQEP FKNLKTGKYA
F2_CM_MP25   R.EILKEPVH GVYYDPSKDL IAEIQKQGHD QWTYQIYQEP HKNLKTGKYA
F2KU_BE_VI   R.EILKEPVH GVYYDPSKDL IAEIQKQGQG QWTYQIYQEP YKNLKTGKYA
G_BE_DRCBL   R.EILKEPVH GVYYDPSKDL IAEVQKQGLG QWTYQVYQEP YKNLKTGKYA
G_NG_92NG0   R.EILKEPVH GVYHDPSKEL IAEVQKQGPD QWTYQIYQEP YKNLKTGKYA
G_SE_SE616   R.EILREPVH GVYYDPSKEL IAEVQKQGLD QWTYQIYQEP YKNLKTGKYA
H_BE_VI991   R.EILKEPVH GAYYDPSKEL IAEIQKQGPD QWTYQIYQEP FKNLKTGKYA
H_BE_VI997   R.EILREPVH GVYYDPSKDL IAEIQKQGPD QWTYQIYQEP FKNLKTGKYA
H_CF_90CF0   R.EILREPIH GVYYDPSKDL IAEIRKQGQG QWTYQIYQEP FKNLKTGKYA
J_SE_SE702   K.EILKEPVH GVYYDPAREL IAEVQKQGLD QWTYQIYQEP FKNLKTGKYA
J_SE_SE788   K.EILKEPVH GVYYDSAKEL IAEVQKQGLD QWTYQIYQEP FKNLKTGKYA
K_CD_EQTB1   R.EILKEPVH GVYYDPSKDL IAEIQKQGHG QWTYQIYQEP YKNLKTGKYA
K_CM_MP535   R.EILKEPVH GVYYDPSKDL IAEIQKQGND QWTYQIYQEP HKNLKTGKYA
N_CM_YBF30   R.EILKEPLH GVYYDPGKEL VAEIQKQGQG QWTYQIYQEL HKNLKTGKYA
O_CM_ANT70   R.ERLKQPVH GVYYQPDKDL WVNIQKQGGE QWTYQIYQEE HKNLKTGKYT
O_CM_MVP51   R.EKLKEPVH GVYYQPDKDL WVSIQKHGEG QWTYQVYQDE HKNLKTGKYA
O_SN_99SE_   R.EKLKEPVH GVYYQPDKDL WVNIQKQGWG QWTYQIYQDE HKNLKTGKYT
O_SN_99SE_   R.EKLKEPVH GVYYQPDKDL WVNIQKQGKG QWTYQIYQDE HKNLKTGKYT
U_CD___83C   R.EILKEPVH GVYYDPSKDL VAEIQKQGQD QWSYQIYQEP FKNLKTGKYA

             551                                                  600
00BW0762_1   KMRTAHTNDV KQLTEAVQKI SRESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW0768_2   KMRTAHTNDV KQLTEAVQKI AQESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW0874_2   KMRTAHTNDV KQLAEAVQKI ATESIVIWG. KTPKFRLPIQ KETWEIWWTD
00BW1471_2   KKRSAHTNDV KQLTDAVQKI AMESIVIWG. KTPKFRLPIQ RETWETWWTN
00BW1616_2   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWDTWWTD
00BW1686_8   KMRTAHTNDV RQLTEAVQKI AMESVVIWG. KTPKFRLPIQ KETWEAWWTD
00BW1759_3   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWAD
00BW1773_2   KMRTAHTNDV KQLTEAVQKI ALESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW1783_5   KMRTAHTNDV KQLTEAVQKI ALESIVIWG. KIPKFRLPIQ KETWETWWTD
00BW1795_6   KRRTAHTNDV KQLTEAVQKI ALESIVIWG. KVPKSRLPIQ KETWETWWTD
00BW1811_3   KMRSAHTNDV KQLTEAVQKI AMECIVIWG. KIPKFRLPIQ KETWETWWTD
00BW1859_5   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW1880_2   KMRTAHTNDV KQLTEVVQKI AMESIVIWG. KAPKFRLPIQ KETWETWWTD
00BW1921_1   KKRSAHTNDV KQLTEAVQKI CIESVVIWG. RTPKFRLPIQ RETWETWWTD
00BW2036_1   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWDTWWTD
00BW2063_6   KRRTAHTNGV KQLTEAVQKI AQESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW2087_2   KMRTAHTNDV KQLTEAVQKI AQESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW2127_2   KMRTAHTNDV KQLTEAVQKI AIESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW2128_3   KMRTAHTNDV KQLTEAVQKI ALESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW2276_7   KMRTAHTNDV KQSTEAVQKI AMESIVIWG. RTPKFRLPIQ KEMWEARWTD
00BW3819_3   KMKSAHTNDV KQLTEAVQKI ALESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW3842_8   KMRTAHTNDV KQLTEAVQKI VMESIVIWG. KAPKFRLPVQ KETWETWWTD
00BW3871_3   KMRNAHTNDV KQLTEAVXKI AMECIVIWG. KTPKFRLPIQ KETWETWWTD
00BW3876_9   RMRSAHTNDV KQLAEAVQRI AKESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW3886_8   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWTD
00BW3891_6   KIRTAHTNDV KQLTEVVQKI AQESIVIWG. KTPKFRLPIQ KETWEIWWTD
00BW3970_2   KRRTAHTNDV RQLAEAVQKI TMESIVIWG. KTPRFRLPVQ KETWETWWTE
00BW5031_1   KIRTAHTNDV RQLTEAVQKI ALESIVIWG. KIPKFRLPIQ KETWETWWTD
 96BW01B21   KMRTAHTNDV KQLTEAVQKI ATESIVIWG. KXPKFRLPIQ KETWETWWTD
  96BW0407   KMRTAHTNDV KQLTEVVQKI AMESIVIWG. KTPKFRLPIQ KDTWETWWTD
  96BW0502   KMRTAHTNDV KQLTEAVQKI AQESIVIWG. KTPKFRLPIQ KETWETWWTD
 96BW06_J4   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KEIWETWWTD
 96BW11_06   KMRTAHTNDV KQLTEAVQKI SMESIVIWG. KIPKFRLPIQ KEAWEAWWTD
  96BW1210   KLRTAHTNDV RQLTEAVQKI AQECIVIWG. KTPKFRLPIQ KETWEAWWTD
 96BW15B03   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWTD
 96BW16_26   KMRTAHTNDV KQLAEAVQKI TMESIVLWG. KTPKFRLPIQ KETWETWWTD
 96BW17A09   KKKSTHTNDV KQLTDAVQKI TMESIVIWG. KTPKFRLPIQ KETWDTWWTE
 96BWMO1_5   KRRTAHTNDV KQLTEAVQKI ALESTVIWG. KIPKFRLPIQ KETWDTWWTD
 96BWMO3_2   KMRTAHTNDV RQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWTD
98BWMC12_2   KMRTAHTNDV KQLAEAVQKI AQESIVIWG. KTPKFRLPIQ KETWETWWTD
```

```
98BWMC13_4   KRKTAHTNDV KQLTEAVQKI ALESIVIWG. KIPKFRLPIQ KETWETWWTD
98BWMC14_a   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWTE
98BWMO14_1   KMRTAHTNDV KQLTEAVQKI AIESIVIWG. KTPKFRLPIQ KETWETWWTD
98BWMO18_d   KMRTTHTNDV RQLTEAVQKI TMESIVIWG. KTPKFRLPIQ KDTWETWWTD
98BWMO36_a   KRRTAHTNDV KQLTEVVQKI TMESIVIWG. KTPKFRLPIQ KDTWETWWTD
98BWMO37_d   KRRTAHTNDV RQLTEAVQKI TMESIVIWG. KTPKFRIPIQ KETWETWWTD
99BW3932_1   KMRSAHPNDV KQLTEAVHKI AMESIVIWG. KTPKFRLPIQ KETWETWWTD
99BW4642_4   KMRTAHTNDV KQLAEAVQKI AVESIVIWG. KTPKFRLPIQ KETWETWWTD
99BW4745_8   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWTD
99BW4754_7   KMRTAHTNDV KQLAEAVQKV AIESIVIWG. KTPKFRLPIQ KETWETWWTD
99BWMC16_8   KIRTAHTNDV RQLTEAVQKI TMECIVIWG. KIPKFRLPIQ KETWETWWTD
A2_CD_97CD   KRKSTHTNDV KQLTEAVQKI ATESIVIWG. KTPKFRLSIQ KDTWETWWTD
A2_CY_94CY   KRRSTHTNDI KQLTEAVQKI TMESIVIWG. KIPKFRLPIQ KETWETWWTE
A2D___97KR   KRRSTHTNDV KQLTEAVQKI VMESIVIWG. KTPKFKLPIQ KETWETWWAE
A2G_CD_97C   KRGSAHTNDV KQLTEVVQKI ATESIVIWG. KIPKFRLPIQ KETWETWWTE
A_BY_97BL0   KXXSAHTNXV KQLTAVVQKV ATESIVIWG. KTPKFKLPIR KETWEVWWTE
A_KE_Q23_A   RKRSAHTNDV RQLAEVVQKV ATESXILXG. KTPKFRLPIQ XETXEXWXMX
A_SE_SE659   RKRSVHTNDV KQLAEMVQKV VMESIVIWG. KTPKFRLPIQ KETWDTWWMD
A_SE_SE725   RKRSAHTNDV RQLAEVVQKV VTESIVIWG. KTPKFKLPIQ KETWETWWMD
A_SE_SE753   RKRSAHTNDV KQLAEVVQKV VMESIVIWG. KTPKFKLPIQ KETWETWWMD
A_SE_SE853   RKRSAHTNDV KQLAEVVQKV AMESIVIWG. KTPRSKLPIQ KETWETWWID
A_SE_SE889   RKRSAHTNDV RQLVEVVQKV VMESIVIWG. KTPKFKVPIQ KETWETWWME
A_SE_UGSE8   RKKSAHTNDV RQLAEVVQKV VMESIVIWG. KTPKFKLPIQ KETWETWWMD
A_UG_92UG0   RKRSAHTNDV KQLAEVVQKV VMESIVIWG. KTPKFKLPIQ KDTWDTWWMD
A_UG_U455_   RKRSAHTNDV KQLTEVVQKV STESIVIWG. KTPKFKLPIQ KETWETWWMD
AC_IN_2130   KMRTAHTNDV KQLTETVQKI AMESIVIWG. KIPKFRLPIQ KETWEAWWME
AC_RW_92RW   KRRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWTD
AC_SE_SE94   KKRSAHTNDV RQLAEVVQKV VMESIIIWG. KTPKFRLPIQ KETWETWWTD
ACD_SE_SE8   KKRSAHTNDV KQLAEVVQKV VMESIVIWG. KTPKFKLPIQ KETWETWWMD
ACG_BE_VI1   KRSAAHTNDV KQLTEAVQKI AQESIVIWG. KAPKFKLPIQ KETWETWWMD
AD_SE_SE69   KVRGTHTNDI KQLTAAVQKI AQESIVIWG. KTPKFRLPIQ RETWEAWWME
AD_SE_SE71   KLRGTHTNDI KQLTAAVQKI AQECIVIWG. KTPKFRLPIQ KETWETWWTE
ADHK_NO_97   RMKSAHTNDV KQLTEAVQKI AQECIVIWG. KTPKFRLPIQ KETWETWWTE
ADK_CD_MAL   RIKSAHTNDV KQLTEAVQKI SIESIVIWG. KTPKFRLPIQ KETWETWWTE
AG_BE_VI11   KRKSAHTNDV KQLTEAVQKI AQESIVIWG. KTPKFRLPIQ KETWEAWWTE
AG_NG_92NG   KRGSAHTNDV KQLVEVVQKV ATESVVIWG. KTPKFRLPIQ RETWEAWWRE
AGHU_GA_VI   KKRSAHTNDV KQLTEAVQKI ATESIVIWG. KVPKFKLPIR KETWEVWWTE
AGU_CD_Z32   RRRTAHTNDV RQLAEVVQKI ATESIVIWG. KIPKFKLPIQ KETWETWWTD
AJ_BW_BW21   KRRAAHTNDV KQLAEVVQKI ASESIVIWG. KTPKFRLPIQ KETWETWWAD
B_AU_VH_AF   RMKGAHTNDV KQLTEAVQKI ALESIVIWG. KTPKFRLPIQ KETWETWWTE
B_CN_RL42_   RMRGAHTNDV KQLTEAVQKI ATESIVIWG. KTPKFKLPIQ KETWEAWWME
B_DE_D31_U   RMRGAHTNDV KQLTEAVQKI ATESIVIWG. KTPKFKLPIQ KETWEAWWTE
B_DE_HAN_U   RMRGAHTNDI KQLTEAVQKI TTESIVIWG. KTPKFKLPIQ KETWEAWWTE
B_FR_HXB2_   RMRGAHTNDV KQLTEAVQKI ATEGIVIWG. KTPKFRLPIQ KETWEAWWTE
B_GA_OYI__   RMRGAHTNDV KQLTEAVQKI TTESIVIWG. KTPKFKLPIQ KETWETWWTE
B_GB_CAM1_   KMRGTHTNDV KQLTEAVQKI TQESIVIWG. KTPKFKLPIQ KETWEAWWTE
B_GB_GB8_A   KMRGAHTNDV KQLTEAVQKI ATESIVIWG. KTPKFKLPIQ KETWDAWWID
B_GB_MANC_   KMRGTHTNDV KQLTEAVQKV ATESIVIWG. KTPKFKLPIQ KETWEAWWME
B_KR_WK_AF   RMRGAHTNDV KQLTEAVQKV TTESIIIWG. KIPKFKLPIQ KETWDAWWTE
B_NL_3202A   RMRGAHTNDV KQLTEAVQKI AIESIVIWG. KTPKFKLPIQ KETWEAWWTE
B_TW_TWCYS   RTRGAHTNDV KQLTEAVQKI TTESIVIWG. KTPKFKLPIQ KETWEAWWTE
B_US_BC_L0   RTRGAHTNDV KQLTEAVQKI ATESIVIWG. KTPKFKLPIQ KETWEAWWTE
B_US_DH123   RTRGAHTNDV KQLTEVVQKV TTECIIIWG. KTPKFRLPIQ KETWETWWTE
B_US_JRCSF   RTRGAHTNDV KQLTEAVQKI TTECIVIWG. KTPKFRLPIQ KETWETWWTE
B_US_MNCG_   RMRGAHTNDV KQLTEAVQKI ANESIVIWG. KIPKFKLPIQ KETWETWWTE
B_US_P896_   RMRGAHTNDV KQLTEAVQKI ATESIVIWG. KTPKFRLPIQ KETWETWWTE
B_US_RF_M1   RMRGAHTNDV KQLTEAVQKV ATESIVIWG. KTPKFKLPIQ KETWEAWWTD
B_US_SF2_K   RMRGAHTNDV KQLTEAVQKV ATESIVIWG. KTPKFKLPIQ KETWEAWWTE
B_US_WEAU1   RVRGAHTNDV KQLTEAVQKV STESIVIWG. KIPKFKLPIQ KETWEAWWME
B_US_WR27_   RTRGAHTNDV KQLTEAVQKI TTESIVIWG. KTPKFKLPIQ KETWETWWTE
B_US_YU2_M   RTRGAHTNDV XQLXEAVQKX ATXSIVIWG. KTPKXILPIQ KETWESRWTE
                         KQLTEAVQKI ATESIVIWG. KTPKFKLPIQ KETWETWWTE
```

```
BF1_BR_93B   RMRGAHTNDV KQLTEAVQKI TTESIVIWG. KIPKFKLPIQ KETWEAWWIE
C_BR_92BR0   KMRTAHTNDV RQLTEAVQKI ALESIIIWG. KTPKFRLPIQ KETWEAWWTD
C_BW_96BW0   KMRTAHTNDV KQLTEVVQKI TMESIVIWG. KTPKFRLPIQ KDTWETWWTD
C_BW_96BW1   KMRTAHTNDV KQLTEAVQKI SMESIVIWG. KIPKFRLPIQ KEAWEAWWTD
C_BW_96BW1   KLRTAHTNDV RQLTEAVQKI AQECIVIWG. KTPKFRLPIQ KETWEAWWTD
C_BW_96BW1   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWEAWWTD
C_ET_ETH22   KRGTAHTNDV KQLTAVVQKI ALESIVIWG. KTPKFRLPIQ KETWETWWTD
C_IN_93IN1   KRRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWEAWWTD
C_IN_93IN9   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWTD
C_IN_93IN9   KMRTAHTNDV KQLAEAVQKI AMESIVIWG. KTPKFRLPIQ KETWETWWTD
C_IN_94IN1   KMRTAHTNDV KQLTEAVQKI AMESIVIWG. KTPKFRLPIQ KETWEAWWTD
C_IN_95IN2   KMRTAHTNDV KQLTEAVQKI ATESIVIWG. KTPKFRLPIQ KETWETWWTD
CRF01_AE_C   RKRSAHTNDV RQLAEVVQKV ATESIVIWG. KTPKFRLPIQ KETWETWWTD
CRF01_AE_C   RKRSAHTNDV RQLAEVVQKV ATESIVIWG. KTPKFRLPIQ RETWETWWAE
CRF01_AE_C   RQRSAHTNDV RQLVEVVQKV STESIVIWG. KTPKFRLPIQ RETWETWWME
CRF01_AE_T   RKRSAHTNDV RQLTEVVQKI ATESIVIWG. KTPKFRLPIQ RETWETWWME
CRF01_AE_T   RKRSAHTNDV RQLTEVVQKI ATESIVIWG. KTPKFRLPIQ RETWETWWME
CRF01_AE_T   RKRSAHTNDV RQLTEVVQKI ATESIVIWG. KTPKFRLPIQ RETWETWWME
CRF01_AE_T   RNRSAHTNDV RQLTEVVQKI ATESIVIWG. KTPKFRLPIQ RETWETWWME
CRF01_AE_T   RRGSAHTNDV RQLTEVVQKV ATESIVIWG. KTPKFRLPIQ KETWETWWME
CRF01_AE_T   RKRSAHTNDV RQLTEVVQKI ATESIVIWG. KTPKFRLPIQ RETWETWWME
CRF02_AG_F   KRRSAHTNDV KQLAEVVQKV VTESIVIWG. KTPKFKLPIQ KETWETWWME
CRF02_AG_F   KRRSAHTNDV KQLTEVVQKV ATESIVIWG. KTPKFRLPIQ RETWEAWWME
CRF02_AG_G   ERVAPYDLSI TELTEVVQKV TTESIIIWG. KTPKFSLPIQ RETWEAWWME
CRF02_AG_N   KKRSAHTNDV KQLTEVVQKV AMESIVIWG. KTPKFRLPIQ RETWEAWWME
CRF02_AG_S   KKRSAHTNDV KQLTEVVQKV ATESIVIWG. KTPKFRLPIQ RETWETWWME
CRF02_AG_S   KKRSAHTNDV KQLTEVVQKV ATESIVVWG. KTPKFRLPIQ RETWEAWWME
CRF03_AB_R   RLRGAHTNDV KQLTEAVQKI ATESIVIWG. KTPKFKLPIQ KETWETWWTE
CRF03_AB_R   RLRSAHTNDV KQLTEAVQKI ATESIVIWG. KTPKFKLPIQ KETWETWWAE
CRF04_cpx_   RTRSAHTNDV RQLTEAVQKI AMECIVIWG. KTPKFRLPIQ KETWDTWWTE
CRF04_cpx_   KTRSAHTNDV RQLTEVVQKI AMECIVIWG. KTPKFRLPIQ KETWDTWWTE
CRF04_cpx_   KTRSAHTNDV RQLTEAVQKI AMECIVIWG. KTPKFRLPIQ KETWDTWWME
CRF05_DF_B   RARGVHTNDV KQLAEAVQKI ATESIVIWG. KTPKFRLPIQ KETWEIWWTE
CRF05_DF_B   RTRNAHTNDV KQLAEAVQKI ATESIVIWG. KTPKFRLPIQ KETWETWWTE
CRF06_cpx_   RIKSAHTNDV KQLTEAVQKI ALESIVIWG. KTPKFRLPIQ KETWETWWTE
CRF06_cpx_   KIKSTHTNDV KQLTEAVQKI ARESIVIWG. KTPKFRLPIQ KETWETWWTE
CRF06_cpx_   RMRSAHTNDI KQLTEAVQKI ALEAIVIWG. KIPKFRLPIQ KETWETWWTE
CRF06_cpx_   RTKSAHTNDV RQLTEAVQKI ALESIVIWG. KTPKFRLPIH KETWETWWTE
CRF11_cpx_   KRRAAHTNDV RQLTEVIQKI CLEAIVIWG. KTPKFRLPIQ RETWETWWTD
CRF11_cpx_   KRRTAHTNDV RQLAEVVQKI ALEGIVIWG. KIPKFRLPIQ KETWETWWTD
D_CD_84ZR0   RMRGAHTNDV KQLTEAVQKI AIESIVIWG. KTPKFRLPIQ KETWETWWID
D_CD_ELI_K   RMRGAHTNDV KQLAEAVQRI STESIVIWG. RTPKFRLPIQ KETWETWWAE
D_CD_NDK_M   RTRGAHTNDV KQLTEAVQKI ATESIVIWG. KTPKFKLPIQ KETWETWWIE
D_UG_94UG1   KMRGTHTNDV KQLTEAVQKI AQECIVIWG. KTPKFRLPIQ KETWETWWTE
F1_BE_VI85   KVRSAHTNDV KQLTEAVQKI ALESIVIWGK RSPKFKLPIL KETWDTWWTD
F1_BR_93BR   KMRSAHTNDV KQLTEAVQKI SLESIVIWG. KTPKFRLPIL KETWDTWWTE
F1_FI_FIN9   KMRSAHTNDV KQLTEAVQKI ALESIVIWG. KTPKFKLPIL KETWDTWWTE
F1_FR_MP41   KTRSAHTNDI KELTDAVQKV ARECIVIWG. KSPKFRLPIQ KETWEAWWTD
F2_CM_MP25   RRKSAHTNDV KQLTEVVQKV ATEGIVIWG. KVPKFRLPIQ KETWEIWWTE
F2KU_BE_VI   RMRSAHTNDV KQLTEAVQKI ATEGIVIWG. KTPKFRLPIQ KETWETWWTE
G_BE_DRCBL   KGGSAHTNDV KQLTEVVQKI ATEGIIIWG. KIPKFKLPIK KETWEVWWTE
G_NG_92NG0   KRGSAHTNDV KQLTEVVQKI ATEGIVIWG. KIPKFKLPIR KETWEVWWTE
G_SE_SE616   KRGSAHTNDV KQLTEVVQKI ATESIVIWG. KTPKFKLPIR KETWEIWWTD
H_BE_VI991   KMRSAHTNDV KQLTEVVQKI ATESIVIWG. KIPKFRLPIQ KETWETWWTE
H_BE_VI997   KMRNAHTSDV KQLTEAVQKI ATESIVIWG. KIPKFKLPIQ KETWETWWTE
H_CF_90CF0   KMRTAHTNDI KQLTEAVQKI STESIVIWG. KIPKFRLPIQ KETWETWWTE
J_SE_SE702   KRRSAHTNDV KQLSQVVQKI ALEAIVIWG. KTPKFRLPIQ KETWETWWTD
J_SE_SE788   KRRSAHTNDV KQLAEVVQKI ALEAIVIWG. KTPKFRLPIQ RETWETWWTD
K_CD_EQTB1   RIRSAHTNDV KQLTEVVQKV AMESIVIWG. KTPKFRLPIQ KETWGTWWTE
K_CM_MP535   RMRSAHTNDV KQLTEAVQKI ATEGIVIWG. KTPKFRLPIQ KETWETWWTE
N_CM_YBF30   KMRSAHTNDI KQLVEVVRKV ATESIVIWG. KTPKFRLPVQ KEVWEAWWTD
```

```
O_CM_ANT70   RQKASHTNDI RQLAEVIQKV SQESIIIWG. KLPKFKLPVT RETWETWWAD
O_CM_MVP51   RQKASHTNDI RQLAEVVQKV SQEAIVIWG. KLPKFRLPVT RETWETWWAE
O_SN_99SE_   RQKASHTNDI RQLAEVLQRV SQEAIIIWG. KLPKFKLPIT RETWETWWAD
O_SN_99SE_   KQKASHTNDI RQLAEVLQKV SQEAIIIWG. KLPKFKLPIT RETWETWWAD
U_CD___83C   KRRSAHTNDV KQLTEAVQKI SMESIVIWG. KIPKFKLPIQ KETWESWWTD

             601                                                  650
00BW0762_1   YWQATWIPEW EFVNTPPLVK LWYQLEKDPL VGVETFYVDG AANRDTKIGK
00BW0768_2   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI LGAETFYVDG AANRETKKGK
00BW0874_2   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AANRETKLGK
00BW1471_2   YWQAT.IPEW EFVNAPPLVK LWYQLEKDPI AGAETFYVDG AANRETKLGK
00BW1616_2   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGIETFYVDG AANRETKLGK
00BW1686_8   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKAGK
00BW1759_3   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AANRDTKKGK
00BW1773_2   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGVETFYVDG AANRETKIGK
00BW1783_5   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETYYVDG AANRETKMGR
00BW1795_6   YWQATWIPEW EFVNTPPLVK LWYQLEKEPV AGVETFYVDG AANRETKMGK
00BW1811_3   YWQATWIPEW EFVNTPPLVK LWYQLEEEPI AGAETFYVDG AANRETKIGK
00BW1859_5   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AAHRETKVEK
00BW1880_2   YWQATWIPEW EFVNTPPLVK LWYQLEKEPM MGVETFYVDG AANRETKIGK
00BW1921_1   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
00BW2036_1   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKLGK
00BW2063_6   YWQATWIPEW EFVNTPPLVK LWYQLEKDPI LGVETFYVDG AANRETKMGK
00BW2087_2   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AANRETKIGK
00BW2127_2   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGVETFYVDG AVNRETKVGK
00BW2128_3   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKKGK
00BW2276_7   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AANRETKIGK
00BW3819_3   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGVETFYVDG AANRDTKVGK
00BW3842_8   YWQATWIPDW EFVNTLPLVK LWYQLEKEPI IGAETFYVDG AANRDTKIGK
00BW3871_3   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKAGK
00BW3876_9   YWQATWIPDW EFVNTPPLVK LWYQLEKEPI VRAETFYVDG AANRETKIGK
00BW3886_8   YWQATWIPEW EFVNTPPLVK LRYQLEKEPI AGVETFYVDG AANRETKVGK
00BW3891_6   YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AEVETFYVDG AANRETKRGK
00BW3970_2   YWQATWIPEW EFVNTPPLVK LWYQLEKEPM AGVETFYVDG AANRETKIGR
00BW5031_1   YWQATWIPEW EFVNTPPLVK LWYQLEKEPV AGVETFYVDG AANRETKIGK
 96BW01B21   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKIGK
  96BW0407   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AANRETKLGK
  96BW0502   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI PGVETFYVDG AANRETKLGK
 96BW06_J4   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI MGAETFYVDG AANRETKAGK
 96BW11_06   Y.QATWIPEW EFVNTPPLVK LWYQLETEPM AGAETFYVDG AANRETKIGK
  96BW1210   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGVETFYVDG AANRETKMGK
 96BW15B03   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKIGK
 96BW16_26   YWQATWIPKW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
 96BW17A09   YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGAETFYVDG AANRETKLGK
 96BWMO1_5   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI LGAETFYVDG AANRETKMGK
 96BWMO3_2   YWQATWIPEW EFVNTPPLVK LWYQLEKGPI AGAETFYVDG ASNRETKLGK
98BWMC12_2   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETLYVDG AANRETKLGK
98BWMC13_4   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI LGVETFYVDG AANRETKLGK
98BWMC14_a   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGVETFYVDG AANRETKAGK
98BWMO14_1   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI PGAETFYVDG AANRETKTGK
98BWMO18_d   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI GGAETFYVDG AANRETKMGK
98BWMO36_a   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AASRETKKGK
98BWMO37_d   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGVETFYVDG AASRDTKIGK
99BW3932_1   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI GGVETFYVDG AANRETKVGK
99BW4642_4   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKKGK
99BW4745_8   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGVETFYVDG ATNRETKRGN
99BW4754_7   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGVETFYVDG AANRDTKLGK
99BWMC16_8   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKKGK
A2_CD_97CD   WWQATWIPEW EFVNTPPLVK LWYQLETEPI AGAETFYVDG AANRETKLGK
A2_CY_94CY   YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AANRETKLGK
A2D___97KR   YWQATWIPEW EFVNTPPLVK LWYQLETEPI TGAETFYVDG AANRETKLGK
```

```
A2G_CD_97C  YWQATWIPDW EFVNTPPLVK LWYRLETEPI PGAETYYVEG AANRETKLGK
A_BY_97BL0  XWQATXIPE. EFVNTPPLVK LWYQLEKEPI VRAETFYVDR AANRETKIGK
A_KE_Q23_A  YWQATWILEW EFVNTPPLVK LWYQLEKDPI AGAETFYVDG AANRETKLGK
A_SE_SE659  YWQATWIPEW EFVNTPPLVL WWYLLEEDPI VGADSFFVEG AAKEKTLXGT
A_SE_SE725  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGAETFYVDG AANRETKLGK
A_SE_SE753  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AANRETKLGK
A_SE_SE853  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGAETFYVDG AANRETKLGK
A_SE_SE889  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AANRETKLGK
A_SE_UGSE8  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGVETFYVDG AANRETKLGK
A_UG_92UG0  YWQAT.IPEW EFVNTPPLVK LWYQLEKDPI AGAETFYVDG AANRETKLGK
A_UG_U455_  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGAETFYVDG AANRETKLGK
AC_IN_2130  YWQATWIPEW EFVNTPPLVK LWYQLEKEPM AGAETFYVDG AANRETKIGK
AC_RW_92RW  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI LGAETFYVDG AANRETKIGK
AC_SE_SE94  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI IGAETFYVDG AANRETKLGK
ACD_SE_SE8  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI IGAETFYVDG AANRETKLGK
ACG_BE_VI1  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AASRETKLGK
AD_SE_SE69  YWQATWIPEW EFVNTPPLVK LWYQLEKDPM VGAETFYVDG AANRETKLGK
AD_SE_SE71  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGVETFYVDG AANRETKLGK
ADHK_NO_97  YWQATWIPEW EFVNTPPLVK LWYQLETEPI VGAETFYVDG AANRETKKGK
ADK_CD_MAL  YWQATWIPEW EFVNTPPLVK LWYQLETEPI VGAETFYVDG AANRETKKGK
AG_BE_VI11  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGAETFYVDG AANRETKLGK
AG_NG_92NG  YWQATWIPDW EFVNTPPLVK LWYRLETEPI PGAETYYVDG AANKETKLGK
AGHU_GA_VI  HWQATWIPEW EFVNTPHLVK LWYQLETEPI VGAETFYVDG AANRETKQGK
AGU_CD_Z32  YWQATWIPEW EFVNTPLLVK LWYQLEKEPI IGAETFYVDG AANRETKQGK
AJ_BW_BW21  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG ASNRETKLGK
B_AU_VH_AF  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
B_CN_RL42_  YWQATWIPEW EFINTSPLVK LWYQLEKEPI EEAETFYVDG AANRETKLGK
B_DE_D31_U  YWQATWIPEW EFVNTPPLVK LWYQLETEPI VGAETFYVDG AANRETKLGK
B_DE_HAN_U  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
B_FR_HXB2_  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
B_GA_OYI__  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AANRETKLGK
B_GB_CAM1_  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETRLGK
B_GB_GB8_A  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKLGK
B_GB_MANC_  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
B_KR_WK_AF  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
B_NL_3202A  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
B_TW_TWCYS  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKLGK
B_US_BC_L0  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI EGAETFYVDG AANRETKLGK
B_US_DH123  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AASRETRLGK
B_US_JRCSF  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
B_US_MNCG_  YT.ATWIPEW EVVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKKGK
B_US_P896_  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRDTKSGK
B_US_RF_M1  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKLGK
B_US_SF2_K  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
B_US_WEAU1  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANKETKLGK
B_US_WR27_  YWQATWIPEW EXVNTPPLVK LWYQLEKEPI VGAETFYVDG AASRETKLGK
B_US_YU2_M  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKLGK
BF1_BR_93B  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGK
C_BR_92BR0  YWQAT.IPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AANREIKMGK
C_BW_96BW0  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AANRETKLGK
C_BW_96BW1  YWQATWIPEW EFVNTPPLVK LWYQLETEPM AGAETFYVDG AANRETKIGK
C_BW_96BW1  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGVETFYVDG AANRETKMGK
C_BW_96BW1  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKIGK
C_ET_ETH22  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGVETFYVDG AANRETKIGK
C_IN_93IN1  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGVETFYVDG AANRETKLGK
C_IN_93IN9  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGVETFYVDG AANRETKKGK
C_IN_93IN9  YWQATWIPEW EFVNTPPLVK LWYRLEKEPI AGVETFYVDG AANRETKIGK
C_IN_94IN1  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGAETFYVDG AANRDTKIGK
C_IN_95IN2  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI AGAETFYVDG AANRETKIGK
CRF01_AE_C  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGAETLYVDG AASRENKLGK
CRF01_AE_C  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI MGAETFYVDG AASRETKQGK
```

```
CRF01_AE_C  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI AGAETFYVDG AASRETKLGK
CRF01_AE_T  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AASRETKLGK
CRF01_AE_T  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AASRETKLGK
CRF01_AE_T  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AASRETKLGK
CRF01_AE_T  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AASRETKLGK
CRF01_AE_T  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AASRETKLGK
CRF01_AE_T  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AASRETKLGK
CRF02_AG_F  YWQATWIPDW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AANRETKLGK
CRF02_AG_F  YWQATWIPDW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AANRETKLGK
CRF02_AG_G  YWQATWIPEW EFVNTLPLVK LWYQLEKDPI VGAETFYVDG AANKETKLGK
CRF02_AG_N  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AANRETKLGK
CRF02_AG_S  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI VGAETFYVDG AANRETKIGK
CRF02_AG_S  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI IGAETFYVDG AANRETKIGK
CRF03_AB_R  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKLGR
CRF03_AB_R  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKSGK
CRF04_cpx_  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKSGK
CRF04_cpx_  YWQATWIPEW EFVNTPPLVK LWYQLETDPI AGAETFYVDG AANRETKQGK
CRF04_cpx_  YWQATWIPEW EFVNTPPLVK LY.QLEPDPI AGAETFYVDG AASRETRRGK
CRF05_DF_B  YWQATWIPEW EFVNTPPLVK LWYQLETEPI AGAETFYVDG AASRETNQGK
CRF05_DF_B  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI LGAETFYVDG AANRETKLGK
CRF06_cpx_  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI MGAETFYVDG AANRETKLGK
CRF06_cpx_  YWQATWIPEW EFVNTPPLVK LWYQLETEPI VGAETFYVDG AANRETKQGK
CRF06_cpx_  YWQATWIPEW EFVNTPPLVK LWYQLETEPI VGAETFYVDG AANRETKKGK
CRF06_cpx_  YWQATWIPEW EFVNTPPLVK LWYQLETEPI VGAETFYVDG AANKETKKGK
CRF06_cpx_  YWQATWIPEW EFVNTPPLVK LWYQLETEPI IGAETFYVDG AANRETKIGK
CRF11_cpx_  YWQATWIPEW EFVNTPPLVK LWYQLEQGPI LGAETFYVDG AANRETKIGK
CRF11_cpx_  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKLGK
D_CD_84ZR0  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANKETKLGK
D_CD_ELI_K  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKLGK
D_CD_NDK_M  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI IGAETFYVDG AANRETKLGK
D_UG_94UG1  YWQATWIPEW EYVNTPPLVK LWYQLEKEPI VGAETFYVDG AANRETKIGK
F1_BE_VI85  YWQATWIPEW EFVNTPPLVK LWYQLETEPI AGADTFYVDG ASNRETKKGK
F1_BR_93BR  YWQATWIPEW EFVNTPPLVK LWYQLETEPI VGAETFYVDG ASNRETKKGK
F1_FI_FIN9  YWQATWIPEW EFVNTPPLVK LWYQLETEPI AGAETFYVDG ASNRETKKGK
F1_FR_MP41  YWQATWIPEW EFVNTPPLVK LWYQLETEPI IGAETFYVDG ASNRETKKGK
F2_CM_MP25  YWQATWIPEW EFVNTPPLVK LWYQLETEPI IGAETFYVDG AANRETKLGK
F2KU_BE_VI  YWQATWIPEW EFVNTPPLVK LWYQLETDPI AGAETFYVDG AANRETKKGR
G_BE_DRCBL  YWQATWIPEW EFVNTPPLVK LWYQLETEPI PGVETYYVDG AANRETKLGK
G_NG_92NG0  YWQAAWIPEW EFVNTPPLVK LWYQLETEPI PGAETYYVDG AANRETKLGK
G_SE_SE616  YWQATWIPEW EFVNTPPLVK LWYRLETEPI PGAETYYVDG AANRETKLGK
H_BE_VI991  HWQATWIPEW EFVNTPHLVK LWYQLETEPI EGAETYYVDG AANRETKMGK
H_BE_VI997  HWQATWIPE. EFVNTPHLVK LWYQLEAEPI AGAETYYVDG AANRETKIGK
H_CF_90CF0  YWQATWIPEW EFVNTPHLVK LWYQLETEPI AGAETYYIDG AANRETKLGK
J_SE_SE702  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI MGAETFYVDG ASNRETKVGK
J_SE_SE788  YWQATWIPEW EFVNTPPLVK LWYQLEKEPI MGAETFYVDG ASNRETKTGK
K_CD_EQTB1  YWQATWIPEW EFVNTPPLVK LWYQLETEPI VGAETFYVDG AANRETKQGK
K_CM_MP535  YWQATWIPEW EFVNTPPLVK LWYQLETEPI VGAETFYVDG AAHRETKKGR
N_CM_YBF30  HWQATWIPEW EFVNTPPLVK LWYQLETEPI SGAETFYVDG AANRETKLGK
O_CM_ANT70  YWQATWIPEW EFVSTPPLIK LWYRLESEPI MGAETYYVDG AANRETKLGK
O_CM_MVP51  YWQATWIPEW EFVSTPPLIK LWYQLETEPI VGAETFYVDG AANRNTKLGK
O_SN_99SE_  YWQATWIPEW EFVSTPPLIK LWYQLESEPI MGAETYYVDG AANRDTKLGK
O_SN_99SE_  YWQATWIPEW EFVSTAPLIK LWYQLESEPI IGAETYYVDG AANRDTKLGK
U_CD___83C  YWQATWIPEW EFVNTPPLVK LWYQLEKDPI MGAETFYVDG AANRETKKGK

            651                                                 700
00BW0762_1  AGYVTDKGRR KVITLTETTN QKTELQAIQL ALQDSGPEVN IVTDSQYALG
00BW0768_2  AGYVTDRGRQ KIVPLTETTN QKTELHAIQL ALQDSGSEVN IVTDSQYALG
00BW0874_2  AGYVTDRGRQ KIVSLTETTN QKTELQAIHL ALQDSGSEVN IVTDSQYAFG
00BW1471_2  AGYVTDKGRQ KVVSLTETTN QKTELQAIKL ALQDSGSEVN IVTDSQYALG
00BW1616_2  AGYVTDRGRQ KIVSLAETTN QKAELQAIQL ALQDSGSEVN IVTDSQYALG
00BW1686_8  AGYVTDRGRQ KVISITETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW1759_3  AGYVTDRGRQ KIVSLTETTN QRTELQAIQL ALQDSGLEVN IVTDSQYALG
```

```
00BW1773_2  AGYVTDRGRQ KIVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW1783_5  AGYVTDRGKQ KVISLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW1795_6  AGYVTDRGRQ KIISLTETTN QKAELHAIQL ALQDSGPEVN IVTDSQYALG
00BW1811_3  AGYVTDKGRQ KTVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW1859_5  AGYVTDRGRQ KIVNLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW1880_2  AGYVTDRGRQ KVVTLTETTN QKTELQAIRL ALQDSGPEVN IVTDSQYALG
00BW1921_1  AGYVTDKGRQ KVITLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW2036_1  AGYVTDRGRQ KIISLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW2063_6  AGYVTDRGRQ KVVSLTGTTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW2087_2  AGYVTDKGRQ KIVCLNETTN QKAELHAIQL ALQDSGLEVN IVTDSQYALG
00BW2127_2  AGYVTDRGRQ KVISINETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW2128_3  AGYVTDKGRQ KVVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW2276_7  AGYVTDRGRQ KIVSLSETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW3819_3  AGYVTDRGRQ KIVSINGTTN QKTELRAIYL ALQDSGSEVN IVTDSQYALG
00BW3842_8  AGYVTDRGRQ KVVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW3871_3  AGYVTDRGRQ KVISLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
00BW3876_9  AGYVTDRGRQ KIVSLTETTN QKSELQAIHL ALQDSGSEVN IVTDSQYALG
00BW3886_8  AGYVTDRGRQ KVITLTETTN QKAELQAIQL ALQDSGSEAN IVTDSQYALG
00BW3891_6  AGYVTDRGRQ KIVSLTETTN QKAELQAIQL ALQDSGSEVN IVTDSQYALG
00BW3970_2  AGYVTDRGRQ KIITLNETTN QKTELQAIKL ALQDSGSEVN IVTDSQYALG
00BW5031_1  AGYVTDRGRQ KIVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
 96BW01B21  AGYVTDRGRQ RIISLTEATN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
  96BW0407  AGYVTNKGRQ KIVPLTETTN QKAELQAIQL ALQDSGSEVN IVTDSQYALG
  96BW0502  AGYVTDRGRQ KIVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
 96BW06_J4  AGYVTDRGRQ KVISLTETTN QKTELQAIQL ALQDSGLEVN IVTDSQYALG
 96BW11_06  AGYVTDKGRQ KVVTLTETTN QKAELQAIQL ALQDSGPEVN IVTDSQYALG
  96BW1210  AGYVTDRGRQ KIVSLNETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
 96BW15B03  AGYVTDRGRQ KIVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
 96BW16_26  AGCVTDRGRQ KIVSLTETTN QKAELQAIQL ALQDSGAEVN IVTDSQYALG
 96BW17A09  AGYVTDKGRQ KVVSLTETTN QKTELQAIKL ALQDSGSEVN IVTDSQYALG
  96BWMO1_5 AGYVTDRGRQ KVVPLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALK
  96BWMO3_2 AGYVTDRGRQ KVVSLNETTN QKTELQAIQL ALQDSGTEVN IVTDSQYALG
98BWMC12_2  AGYVTDKGRQ KVVSLTETTN QKTELHAIQL ALQDSGSEVN IVTDSQYALG
98BWMC13_4  AGYITDRGRQ KVVSLTETTN QKTELQAIQL ALQDSGPEVN IVTDSQYALG
98BWMC14_a  AGYVTDRGRQ KVISLTETAN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
98BWMO14_1  AGYVTDKGRQ KIVSLNETTN QKAELQAIQL ALQDSGLEVN IVTDSQYALG
98BWMO18_d  AGYVTDKGRQ KIVALTETTN QKAELQAIQL ALQDSGSEVN IVTDSQYALG
98BWMO36_a  AGYVTDRGRQ KCVTLTETTN QKTELQAIQL ALQDSGPEVN IVTDSQYALG
98BWMO37_d  AGYVTDRGKQ KIITLTETTN QKTELQAIQL ALQDSGPEVN IVTDSQYALG
99BW3932_1  AGYVTDKGRQ KVITINETTN QKTELQAIYL ALQDSGSEVN IVTDSQYALG
99BW4642_4  AGYVTDRGRQ RIVNLTETTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
99BW4745_8  AGYVTDRGRQ KIVSLTETTN QKAELQAIQL ALQDSGSEVN IVTDSQYALG
99BW4754_7  AGYVTDRGRQ KVVSLTETTN QKTELQAIHL ALQDSGPEVN IVTDSQYALG
99BWMC16_8  AGYVTDRGRQ KVVTLTETTN QKTELQAIQL ALQDSESEVN IVTDSQYALG
A2_CD_97CD  AGYVTDRGRQ KIVPLTETTN QKTELHAIYL ALQDSGSEVN IVTDSQYALG
A2_CY_94CY  AGYVTDRGRQ KIVSLTETTN QKTELHAIYL ALQDSGLEVN IVTDSQYALG
A2D___97KR  AGYVTDKGRQ KIVSLTDTTN QKTELHAIYL ALQDSGLEVN IVTDSQYALG
A2G_CD_97C  AGYVTDKGKQ KIINLTETTN QKTELHAIQL ALQDSGSEVN IVTDSQYALG
A_BY_97BL0  AGYVTDR.RQ KVVPLTETTN QKTELHAIHL VLQDSGSEVD IVTDSQYALG
A_KE_Q23_A  AGYVTDKGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
A_SE_SE659  GWFVAEGGRQ VVVPLAKTTH QTTELPALHP SLQDSGSVVI IVNDSQSAGG
A_SE_SE725  AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
A_SE_SE753  AGYVTNRGRQ KVVSLTETTN QKTELHAILL ALQDSGSEVN IVTDSQYALG
A_SE_SE853  AGYVTDRGRQ KVVSLTETTN QKTEVHAIHL ALQDSGSEVN IVTDSQYALG
A_SE_SE889  AGYVTDRGRQ KVVTLAETTN QKTELHAIYL ALQDSGSEVN IVTDSQYALG
A_SE_UGSE8  AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
A_UG_92UG0  AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
A_UG_U455_  AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
AC_IN_2130  AGYATDRGRQ KIVTLTETTN QKTELQAIYL ALQDSGSEVN IVTDSQYALG
AC_RW_92RW  AGYVTDRGRQ KIVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
AC_SE_SE94  AGYVTDRGRQ KVVSLTETTN QKTELHAIQL ALQDSGSEVN IVTDSQYALG
```

```
ACD_SE_SE8   AGYVTDRGRQ KVVSLTETTN QKTELHAIYL ALQDSGSEVN IVTDSQYALG
ACG_BE_VI1   AGYVTDRGRQ KAITLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
AD_SE_SE69   AGYVTDRGRQ KVVPLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
AD_SE_SE71   AGYITDRGRQ KVVSLTDTTN QKTELQAINL ALQDSGPEVN IVTDSQYALG
ADHK_NO_97   AGYVTDRGRQ KVVFLTETTN QKTELQAIHL ALQDSGSEVN IVTDSQYALG
ADK_CD_MAL   AGYVTDRGRQ KVVSLTETTN QKTELQAIHL ALQDSGSEVN IVTDSQYALG
AG_BE_VI11   AGYVTDRGRQ KAVSLTETTN QKTELHAIQL ALQDSGSEVN IVTDSQYALG
AG_NG_92NG   AGYVTDRGKQ KIITIQETTN QKTELHAIQL ALQDSGSEVN IVTDSQYALG
AGHU_GA_VI   AGYVTDRGRQ KIVSLTETTN QRTELQAIHL ALQDSGPEVN IVTDSQYALG
AGU_CD_Z32   AGYITDKGRQ KVITLTETTN QKTELEAIHL ALQDSGLEVN IVTDSQYALG
AJ_BW_BW21   AGYVTDRGRQ KIVSLTETTN QKTELHAIYL ALQDSGSEVN IVTDSQYALG
B_AU_VH_AF   AGYVTDRGRQ KVVPLSDTTN QKTELQAIQL ALQDSGLEVN IVTDSQYALG
B_CN_RL42_   AGYVTNKGRQ KVVTLTDTTN QKTELQAIYL ALQDSGVEVN IVTDSQYALG
B_DE_D31_U   AGYVTDRGRQ KVVSLTDTTN QKTELQAIHL ALQDSGLEVN IVSDSQYAIG
B_DE_HAN_U   AGYVTDRGRQ KVVSLNDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_FR_HXB2_   AGYVTNRGRQ KVVTLTDTTN QKTELQAIYL ALQDSGLEVN IVTDSQYALG
B_GA_OYI__   AGYVTDRGRQ KVVSLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_GB_CAM1_   AGYVTDRGRQ KVVPLTDTTN QKTELQAIYL ALQDSGLEVN IVTDSQYALG
B_GB_GB8_A   AGYVTDRGRQ KVVPLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_GB_MANC_   AGYVTDRGRQ KVISLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_KR_WK_AF   AGYVTDRGRQ KVVPLTDTTN QKTELQAIHL ALQDSGLEVN IVSDSQYALG
B_NL_3202A   AGYVTDRGRQ KVVSLNDTTN QKTELQAINL ALQDSGLEVN IVTDSQYALG
B_TW_TWCYS   AGYVTDKGRQ KVVSLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_US_BC_L0   AGYVTNKGRQ KVVTLTDTTN QKTELEAIHL ALQDSGLEVN IVTDSQYALG
B_US_DH123   AGYVTNRGRQ KVVSLTDTTN QKTELQAIYL ALQDSGLEVN IVTDSQYALG
B_US_JRCSF   AGYVTSRGRQ KVVSLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_US_MNCG_   AGYVTNRGRQ KVVSLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_US_P896_   AGYVTDRGRQ KVVSLADTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_US_RF_M1   AGYVTDRGRQ KVVSLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_US_SF2_K   AGYVTDRGRQ KVVSIADTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_US_WEAU1   AGYVTNRGRQ RVVSLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_US_WR27_   AGYVTDRGRQ KVVSLNDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
B_US_YU2_M   AGYVTNKGRQ KVVSLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
BF1_BR_93B   AGYVTDRGRQ KVVPLTDTTN QKTELQAIYL ALQDSGLEVN IVTDSQYALG
C_BR_92BR0   AGYVTDRGRQ KIVSITETTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
C_BW_96BW0   AGYVTDKGRQ KIVPLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
C_BW_96BW1   AGYVTDKGRQ EVVTLTETTN QRAELQAIQL ALQDSGSEVN IVTDSQYALG
C_BW_96BW1   AGYVTDRGRQ KIVSLTETTN QKAELQAIQL ALQDSGPEVN IVTDSQYALG
C_BW_96BW1   AGYVTDRGRQ KIVSLNETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
C_BW_96BW1   AGYVTDRGRQ KIVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
C_ET_ETH22   AGYVTDRGRQ KIVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
C_IN_93IN1   AGYVTDRGRQ KIVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
C_IN_93IN9   AGYVTDRGRQ KIVSLTETTN QKTELQAIYL ALQDSGSEVN IVTDSQYALG
C_IN_93IN9   AGYVTDRGRQ KIVSLTETTN QKTELQAICL ALQDSGSEVN IVTDSQYALG
C_IN_94IN1   AGYVTDRGRQ KIVPLTETTN QKTELQAIYL ALQDSGSEVN IVTDSQYALG
C_IN_95IN2   AGYVTDRGRQ KIVSLTETTN QKTELQAICL ALQDSGSEVN IVTDSQYALG
CRF01_AE_C   AGYVTDKGRN KIVSLTETTN QKTELQAICL ALQDSGSEVN IVTDSQYALG
CRF01_AE_C   AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSDVN IVTDSQHALG
CRF01_AE_C   AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGPEVN IVTDSQYALG
CRF01_AE_C   AGYATDRGRH KVVSLTETTN QKTELHAIRL ALRDSGSEVN IVTDSQYALG
CRF01_AE_T   AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
CRF01_AE_T   AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
CRF01_AE_T   AGYVTDRGRQ KVISLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
CRF01_AE_T   AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
CRF01_AE_T   AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
CRF01_AE_T   AGYVTNRGRQ KVVSLTETTN QKSELHAIHL ALQDSGSEVN IVTDSQYALG
CRF02_AG_F   AGYVTDRGRQ KVVSLTETTN QKTELHAIYL ALQDSGSEVN IVTDSQYALG
CRF02_AG_F   AGYVTDRGRQ KVVSLTETTN QKTELHAIYL ALQDSGSEVN IVTDSQYALG
CRF02_AG_G   AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
CRF02_AG_N   AGYVTDRGRQ KVVSLTETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
CRF02_AG_S   AGYVTDRGRQ KVVSLAETTN QKTELHAIHL ALQDSGSEVN IVTDSQYALG
CRF02_AG_S   AGYVTDRGRQ NVVSLTETTN QKTELHAILL ALQESGSEVN IVTDSQYALG
CRF03_AB_R   AGYVTDRGRQ KVVSLTDTTN QKTELQAIHL ALQDSGLEVN IVTDSQYALG
```

```
CRF03_AB_R  AGYVTDRGRQ KVVSLTDITN QKTELHAIHL ALQDSGLEVN IVTDSQYALG
CRF04_cpx_  AGYVTDRGRQ KVVSLSETTN QKTELQAIYL ALQDSGSEVN IVTDSQYAIG
CRF04_cpx_  AGYVTDRGRQ KVVSLSETTN QRTELQAIYL ALKDSGSEVN IVTDSQYAIG
CRF04_cpx_  AGYVTDRGRQ KVITLPETTN QKTELQAIYL ALQDSGSEVN IVTDSQYAMG
CRF05_DF_B  AGYITDKGRQ KAVSLTETTN QKAELQAIHL ALQDSGLEVN IVTDSQYALG
CRF05_DF_B  AGYVTDKGRQ KAVSLTETTN QKTELHAIYL ALQDSGSEVN IVTDSQYALG
CRF06_cpx_  AGYVTDRGRQ KVVTLTETTN QKTELQAINL ALQDSGPEVN IVTDSQYALG
CRF06_cpx_  AGYVTDRGRQ KVVSLAETTN QKTELQAINL ALQDSGSEVN IVTDSQYALG
CRF06_cpx_  AGYVTDKGRQ RVISLTETTN QKTELQAINL ALQDSGSEVN IVTDSQYALG
CRF06_cpx_  AGYVTDRGRQ KVVSLTETTN QKTELQAINL ALQDSGSEVN IVTDSQYALG
CRF11_cpx_  AGYVTDKGRQ KVITLTETTN QKTELEAIHL ALQDSGLEVN IVTDSQYALG
CRF11_cpx_  AGYVTNKGRQ KVITPIETTN QKTELEAIHL ALKDSGLEVN IVTDSQYALG
D_CD_84ZR0  AGYVTDRGRQ KVVPFTDTTN QKTELQAINL ALQDSGLEVN IVTDSQYALG
D_CD_ELI_K  AGYVTDRGRQ KVVPLTDTTN QKTELQAINL ALQDSGLEVN IVTDSQYALG
D_CD_NDK_M  AGYVTDRGRQ KVVPFTDTTN QKTELQAINL ALQDSGLEVN IVTDSQYALG
D_UG_94UG1  AGYVTDRGRQ KVVSLTDTTN QKTELQAINL ALQDSGLEVN IVTDSQYALG
F1_BE_VI85  AGYVTDKGKQ KVVSLTETTN QKAELQAIYL ALQDSGSEVN IVTDSQYALG
F1_BR_93BR  AGYVTDRGRQ KAVSLTETTN QKAELQAIQL ALQDSGSEVN IVTDSQYALG
F1_FI_FIN9  AGYVTDRGRQ KVVSLTETTN QKAELQAIHL ALQDSGSEVN IVTDSQYALG
F1_FR_MP41  AGYVTDRGRQ KAVILTETTN QKAELQAIHL ALQDSGSEVN IVTDSQYVLG
F2_CM_MP25  AGYITDRGRQ KVVSLTETTN QKTELQAIQL ALQDSGSEVN IVTDSQYALG
F2KU_BE_VI  AGYVTDKGRQ KVVSLTETTN QKTELQTIYL ALQDSGSEVN IVTDSQYAIG
G_BE_DRCBL  AGYVTDKGRQ IIITLTETTN QKAELHAIQL ALQDSQSEVN IVTDSQYALG
G_NG_92NG0  AGHVTDKGKQ KIITLTETTN QKAELHAIQL ALQDSRPEVN IVTDSQYALG
G_SE_SE616  AGYVTDKGKQ KIITLTETTN QKAELQAIQL ALQDSRSEVN IVTDSQYALG
H_BE_VI991  AGYVTDRGKQ KIVSLTETTN QKTELQAIYL ALQESGPEVN IVTDSQYALG
H_BE_VI997  AGYVTDKGKQ KVVALTETTN QKTELQAIYL ALQDSGLEVN IVTDSQYALG
H_CF_90CF0  AGYVTDRGKQ KVVSLTETTN QKTELQAIYL ALQDSGLEVN IVTDSQYALG
J_SE_SE702  AGYVTDKGRQ KVITLTDTTN QKTELQAIYL ALQDSGIEVN IVTDSQYALG
J_SE_SE788  AGYVTDKGRQ KVVTLTDTTN QKTELHAIYL ALRDSGLEVN IVTDSQYALG
K_CD_EQTB1  AGYVTDKGRQ KVISITETTN QKTELQAIHL ALQDSGSEVN IVTDSQYALG
K_CM_MP535  AGYVTDRGRQ KVVSITETTN QKAELQAICL ALQDSGSEVN IVTDSQYALG
N_CM_YBF30  AGFVTDRGRQ KVVSIADTTN QKAELQAILM ALQESGRDVN IVTDSQYAMG
O_CM_ANT70  AGYVTEQGKQ KIIKLDETTN QKAELMAILL ALQDSKETVN IVTDSQYALG
O_CM_MVP51  AGYVTEQGKQ NIIKLEETTN QKAELMAVLI ALQDSKEQVN IVTDSQYVLG
O_SN_99SE_  AGYVTEKGRQ KIIKLEETTN QRAELMAVLL ALQDSKETVN IVQDSQYVLG
O_SN_99SE_  AGYVTEKGKQ KIVKLEETTN QKAELMAVLL ALQDSKETVN IVTDSQYVLG
U_CD___83C  AGYVTDRGRQ KVISLTETTN QKAELQALYL ALQDSGLEAN IVTDSQYALG

            701                                                 750
00BW0762_1  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQIDKLVSS
00BW0768_2  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSK
00BW0874_2  IIQAQPDKSE SELVNQIIEQ LINKERVYLS WVPAHKGIGG NEQVDKLVSS
00BW1471_2  IIQAQPDDSE SELVNQIIEQ LIKKDKVYLS WVPAHKGIGG NEQVDKLVSS
00BW1616_2  IIQAQPDKSE SELVNQIIEQ LINKERIYLS WVPAHKGIGG NEQVDKLVSS
00BW1686_8  IIQAQPDKSE SELVNQIIEQ LINKERVYLS WVPAHKGIGG NEQVDKLVSN
00BW1759_3  IIQAQPDKSE SELVNQIIEE LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
00BW1773_2  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSK
00BW1783_5  IIQAQPDKSE SELVNQIIEQ LINKERVYLS WVPAHKGIGG NEQVDKLVSN
00BW1795_6  IIQAQPDKSE SELVSQIIEQ LINKERVYLS WVPAHKGIGG NEQVDKLVSS
00BW1811_3  IIQAQPDKSE SELVNQIIEQ LIRKERVYLS WVPAHKGIGG NERVDKLVSS
00BW1859_5  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGE NEQVDKLVSS
00BW1880_2  ITQAQPDKSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSN
00BW1921_1  IIQAQPDKSE SELVNQIIEE LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
00BW2036_1  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSR
00BW2063_6  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
00BW2087_2  IIQAQPDKSE SELVNQIIEQ LINKEKIYLS WVPAHKGIGG NEQVDKLVSN
00BW2127_2  IIQAQPDKSE SELVNQIIEQ LIKKDRVYLS WVPAHKGIGG NEQVDKLVSS
00BW2128_3  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSN
00BW2276_7  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSK
00BW3819_3  IIQAQPDKSE SEIVNQIIEQ LIQKDRVYLS WVPAHKGIGG NEQVDKLVSS
```

```
00BW3842_8  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
00BW3871_3  IIQAQPDKSE SELVNQIIEE LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
00BW3876_9  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
00BW3886_8  IIQAQPDKSE SELVNQIIEQ LIKKEGVYLS WVPAHKGIGG NEQVDKLVSK
00BW3891_6  IIQSQPGKSE SELVNQKIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSK
00BW3970_2  IIQAQPDTSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
00BW5031_1  IIQAQPDKSE SELVNQIIEQ LIKKERIYLS WVPAHKGIGG NEQIDKLVSS
 96BW01B21  IIQAQPDKSE SELVNQMIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
  96BW0407  IIQAQPDKSE SELVTQIIEQ LIKKERIYLS WVPAHKGIGG NEQVDKLVSS
  96BW0502  IIQAQPDKSE SELVNQIIEQ LIQKEWVYLS WVPAHKGIGG NEQVDKLVSQ
 96BW06_J4  IIQAQPDKSE SELVNQIIEQ LINKERVYLS WVPAHKGIGG NEQVDKLVSN
 96BW11_06  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEKVDKLVSS
  96BW1210  IIQAQPDKSE SELVNQIIEQ LICKERVYLS WVPAHKGIGG NEQVDKLVSS
 96BW15B03  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQIDKLVSK
 96BW16_26  IIQAQPDKSE SELVNQIIEQ LINKERIYLS WVPAHKGIGG HEQVDKLVSS
 96BW17A09  IIQAQPDESE SELVNQIIEQ LIQKDKVYLS WVPAHKGIGG NEQVDKLVSS
 96BWMO1_5  IIQAQPDKSE SELVNQIIEQ LIKKDRVYLS WVPAHKGIGG NEQVDKLVSS
 96BWMO3_2  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
98BWMC12_2  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
98BWMC13_4  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
98BWMC14_a  IIQAQPDKSE SELVNQIIEQ LIQKERVYLS WVPAHKGIGG NEQVDKLVSS
98BWMO14_1  IIQAQPDKSE SELVNQIIEQ LINKEKIYLS WVPAHKGIGG NEQVDKLVSS
98BWMO18_d  IIQAQPDKSE SEIVNQIIEQ LINKERVYLS WVPAHKGIGG NEQVDKLVSR
98BWMO36_a  IIQAQPDKSE SELVNQIIEQ LINKEKIYLS WVPAHKGIGG NEQVDKLVSS
98BWMO37_d  IIQAQPDNSE SELVNQIIEQ LIQKERVYLS WVPAHKGIGG NEQVDKLVSK
99BW3932_1  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
99BW4642_4  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSN
99BW4745_8  IIQAQPDKSE AELVNQIIEQ LIQKERVYLS WVPAHKGIGG NEQVDKLVSK
99BW4754_7  IIQAQPDKSE SELVNQIIEQ LIKKKKVYLS WVPAHKGIGG NEQVDKLVSS
99BWMC16_8  IIQAQPDKSE SELVNQIIGE LIKKERVYLS WVPAHKGIGG NEQVDKLVSN
A2_CD_97CD  IIQAQPDSSE SELVNQIIEK LIEKERVYLS WVPAHKGIGG NEQVDKLVSC
A2_CY_94CY  IIQAQPERSE SEIVNQIIEK LIEKERVYLS WVPAHKGIGG NEQVDKLVSS
A2D___97KR  IIQAQPDRSE SEIVNLIIEK LIEKERVYLS WVPAHKGIGE NEQVDKLVSS
A2G_CD_97C  IIQAQPDSSE AEIVNQIIEQ LIRKEKVYLS WVPAHKGIGG NEQVDKLVSS
A_BY_97BL0  IISAQPDRSE SEIVNKIIEK LIKKEXVYLX .VPAHKRIXX NEQVDKLVSN
A_KE_Q23_A  IIQAQPDKSE SEIVNQIIEK LIEKDKVYLS WVPAHKGIGG NEQVDKLVSS
A_SE_SE659  IIQAQPDRSE SELVNQIIEK LVGKDKVYLA WVPAHKGIGG NEQVDKLVSS
A_SE_SE725  IIQAQPDSSE SEIVNQIIEK LIGKDKVYLS WVPAHKGIGG NEQVDKLVSS
A_SE_SE753  IIQAQPDRSE SELVNQIIEK LIGKDKVYLS WVPAHKGIGG NEQVDKLVSS
A_SE_SE853  IIQAQPDSSE SELVNQIIEK LIGKDKVYLS WVPAHKGIGG NEQVDKLVSS
A_SE_SE889  IIQAQPDRSE SELVNQIIEK LIGKDKVYLS WVPAHKGIGG NEQVDKLVSA
A_SE_UGSE8  IIQAQPDKSE SELVNQIIEK LIEKDKVYLS WVPAHKGIGG NEQVDKLVSS
A_UG_92UG0  IIQAQPDRSE SELVNQIIEK LIEKDKVYLS WVPAHKGIGG NEQVDKLVSS
A_UG_U455_  IIQAQPDRSE SEIVNQIIEK LIEKEKVYLS WVPAHKGIGG NEQVDKLVSS
AC_IN_2130  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WIPAHKGIGG NEQVDKLVSN
AC_RW_92RW  IIQAQPDSSE SEAVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
AC_SE_SE94  IIQAQPDRSE SEIVNQIIEK LIQKDKVYLS WVPAHKGIGG NEQVDKLVSS
ACD_SE_SE8  IIQAQPDRSE SELVNQIIEK LIGKDKVYLS WVPAHKGIGG NEQVDKLVSS
ACG_BE_VI1  IIQAQPDKSE SELVNQIIEQ LIRKDRVYLS WVPAHKGIGG NEQVDKLVSS
AD_SE_SE69  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLS WVPAHRGIGG NEQVDKLVSN
AD_SE_SE71  IIQAQPDRSE SELVNQIIEK LIGKDKVYLS WVPAHKGIGG NEQVDKLVSS
ADHK_NO_97  IIQAQPDKSE SDLVNQIIEQ LIQKDKVYLT WVPAHKGIGG NEQVDKLVSS
ADK_CD_MAL  IIQAQPDKSE SEIVNQIIEQ LIQKDKVYLS WVPAHKGIGG NEQVDKLVSS
AG_BE_VI11  IIQAQPDRSE SELVNQIIEK LIEKDKVYVS WVPAHKGIGG NEQVDKLVSS
AG_NG_92NG  IIQAQPDRSE SELVNQIIEQ LIKKEKVYLT WVPAHKGIGG NEQVDKLVSS
AGHU_GA_VI  IIQAQPDKSE SELVNQIIEK LIQKNKVYLS WVPAHKGIGG NEKVDKLVSA
AGU_CD_Z32  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
AJ_BW_BW21  IIQAQPDKSE SELVSQIIEE LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
B_AU_VH_AF  IIQAQPDKSE SELVNQIIEQ LIKKEKIYLA WVPAHKGIGG NDKVDSLVSA
B_CN_RL42_  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQIDKLVSA
B_DE_D31_U  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLT WVPAHKGIGG NEQVDKLVSA
```

```
B_DE_HAN_U  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_FR_HXB2_  IIQAQPDQSE SELVNQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_GA_OYI__  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_GB_CAM1_  IIQAQPDKSE SELVSQIIEE LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_GB_GB8_A  IIQAQPDRSE SEVVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_GB_MANC_  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSS
B_KR_WK_AF  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_NL_3202A  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_TW_TWCYS  IIQAQPDKSE SELVSQIIEE LIKKEKVYLT WVPAHKGIGG NEQVDKLVSA
B_US_BC_L0  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_US_DH123  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_US_JRCSF  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVST
B_US_MNCG_  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_US_P896_  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_US_RF_M1  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_US_SF2_K  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDRLVST
B_US_WEAU1  IIQAQPDQSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
B_US_WR27_  IIQAQPDKSE SELVSQIIEQ LIXKXXVYLA WVPAHKGIGG NEQVDKLVSS
B_US_YU2_M  IIQAQPDRSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSS
BF1_BR_93B  IIQAQPDKSE LEIVNQIIEQ LIKKEKIYLA WVPAHKGIGG NEQVDKLVSA
C_BR_92BR0  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
C_BW_96BW0  IIQAQPDKSE SELVTQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
C_BW_96BW1  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
C_BW_96BW1  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEKVDKLVSS
C_BW_96BW1  IIQAQPDKSE SELVNQIIEQ LICKERVYLS WVPAHKGIGG NEQVDKLVSS
C_ET_ETH22  IILAQPDKSE SEIVNQIIEQ LISKERVYLS WVPAHKGIGG NEQIDKLVSK
C_IN_93IN1  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
C_IN_93IN9  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
C_IN_93IN9  IIQAQPDKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
C_IN_94IN1  IIQAQPDKSE SELVNQIIEQ LINKERVYLS WVPAHKGIGG NEQVDKLVSS
C_IN_95IN2  IIQAQPNKSE SELVNQIIEQ LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
CRF01_AE_C  IIQAQPDRSE SEIVNQIIEK LIEKERVYLS WVPAHKGIGG NEQVDKLVSS
CRF01_AE_C  IIQAQPDRSE SDIVNQIIEK LIEKERVYLS WVPAHKGIGG NEQVDKLVSS
CRF01_AE_C  IIQAQPDRSE SEIVNQIIEK LIEKEKVYLS WVPAHKGIGG NEQVDKLVSS
CRF01_AE_T  IIQAQPDRSE SEVVNQIIEE LIKKEKVYLS RVPAHKGIGG NEQVDKLVSS
CRF01_AE_T  IIQAQPDRSE SEVVSQMIEE LIKKEKDYLS WVPAHKGIGG DEQVDKLVSS
CRF01_AE_T  IIQAQPDRSE SEVVSQIIEE LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
CRF01_AE_T  IIQAQPDRSE SEVVNQLIEE LIKKEKVYLS WVPAHKGIGG NEQVDKLVIS
CRF01_AE_T  IIQAQPDRSE SEVVNQIIEE LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
CRF01_AE_T  IIQAQPDRSE SEVVNQIIEE LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
CRF02_AG_F  IIQAQPDRSE SELVNQIIEK LIEKDKVYLS WVPAHKGIGG NEQVDKLVSS
CRF02_AG_F  IIQAQPDRSE SELVNQIIEK LIEKDKVYLS WVPAHKGIGG NEQVDKLVSN
CRF02_AG_G  IIQAQPDRSE SELVNQIIEK LIEKDKVYLS WVPAHKGIGG NEQVDKLVSN
CRF02_AG_N  IIQAQPDRSE SELVNQIIEK LIGKDKVYLS WVPAHKGIGG NEQVDNLVSS
CRF02_AG_S  IIQAQPDRSE SELVTQIIEE LIKKDRVYLS WVPAHKGIGG NEQVDKLVSN
CRF02_AG_S  IIQAQPDRSE SELVNQIIEK LIEKDKVYLS WVPAHKGIGG NEQVDKLVSS
CRF03_AB_R  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSK
CRF03_AB_R  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSA
CRF04_cpx_  IIQAQPDRSE SDLVNQIIEQ LIRKDKVYLS WVPAHKGIGG NEQVDKLVSA
CRF04_cpx_  IIQAQPDRSE SDLVNQIIEL LIQKDKVYLS WVPAHKGIGG NEQVDKLVSN
CRF04_cpx_  IIQAKPDKSE SDLVNQIIEQ LIQKDKVYLS WVPAHKGIGG NEQVDKLVSN
CRF05_DF_B  IIQAQPDKSE SELVNQIIEQ LIEKKKVYLS WVPAHKGIGG NEQVDKLVSN
CRF05_DF_B  IIQAQPDKSE SELVNKIIEQ LVQKERVYLS WVPAHKGIGG NEQVDKLVSA
CRF06_cpx_  IIQAQPDKSE SELVNQIIEQ LIKKEKIYLS WVPAHKGIGG NEQVDKLVSN
CRF06_cpx_  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSN
CRF06_cpx_  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
CRF06_cpx_  IIQAQPDRSE SELVNQIIEQ LINKEKVYLS WVPAHKGIGG NEQVDKLVSS
CRF11_cpx_  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLT WVPAHKGIGG NEQIDKLVSS
CRF11_cpx_  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
D_CD_84ZR0  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSS
D_CD_ELI_K  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSQ
```

```
D_CD_NDK_M  IIQAQPDKSE SELVSQIIEQ LIKKEKVYLA WVPAHKGIGG NEQVDKLVSQ
D_UG_94UG1  IIQAQPDKSE SEVVSQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSN
F1_BE_VI85  IIQAQPDKSE SEIVNQIIEQ LIQKERVYLS WVPAHKGIGG NEQVDKLVSA
F1_BR_93BR  IIQAQPDKSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSA
F1_FI_FIN9  IIQAQPDKSE SELVNQIIEQ LIQKEKIYLS WVPAHKGIGG NEQVDKLVSA
F1_FR_MP41  IIQAQPDKSE SEIVNKIIEK LIQKERVYLS WVPAHKGIGG NEQVDRLVSA
F2_CM_MP25  IIQAHPDKSE SEIVNQIIEQ LIQKERVYLS WVPAHKGIGG NEQVDKLVST
F2KU_BE_VI  IIQAQPDKSE SELVNQIIEQ LIQKDQIYLS WVPAHKGIGG NEQVDKLVSS
G_BE_DRCBL  IIQAQPDRSE SEIVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
G_NG_92NG0  IIQAQPDRSG SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
G_SE_SE616  IIQAQPDRSE AELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
H_BE_VI991  IIQAQPDKSE SELVNQIIEE LIKKEKFYLS WVPAHKGIGG NEQVDKLVSS
H_BE_VI997  IIQAQPDKSE SELVNQIIEE LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
H_CF_90CF0  IIQAQPDKSE SELVNQIIEE LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
J_SE_SE702  IIQAQPDKSE SELVNQIIEE LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
J_SE_SE788  IIQAQPDKSE SELVNQIIEE LIKKEKVYLS WVPAHKGIGG NEQVDKLVSS
K_CD_EQTB1  IIQAQPDKSE SELVNQIIEQ LIKKDRVYLS WVPAHKGIGG NEQVDKLVSS
K_CM_MP535  IIQAQPDKSE SDLVNQIIEQ LIKKERIYLS WVPAHKGIGG NEQVDKLVSA
N_CM_YBF30  IIHSQPDKSE SELVSQIIEE LIKKERVYLS WVPAHKGIGG NEQVDKLVSS
O_CM_ANT70  VISSQPTQSE SPIVQQIIEE LTKKEQVYLT WVPAHKGIGG NEKIDKLVSK
O_CM_MVP51  IISSQPTQSD SPIVQQIIEE LTKKERVYLT WVPAHKGIGG NEKIDKLVSK
O_SN_99SE_  IISSQPTQSE SSLVQQIIEE LTKKEQVYLT WVPAHKGIGG NEKIDKLVSK
O_SN_99SE_  IISSQPTQSE SPLVQQIIEE LTQKEQVFLT WVPAHKGIGG NEKIDKLVSK
U_CD___83C  IIQAQPDSSE SELVNQIIEQ LIKKEKVYLS WVPAHKGIGG NEQADKLVSA

            751                                                 800
00BW0762_1  GIRKVLFLDG IDKAQDEHEK YHSNWRAMAN EFNLPPIVAK EIVASCDKCQ
00BW0768_2  GIRRVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW0874_2  GIRKALFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW1471_2  GIRKVLFLDG IDKAQEEHER YHSNWRAMAN EFNLPPIVAK EIVASCDKCQ
00BW1616_2  GIRKVLFLDG IDKAQEEHEK YHSSWRAMAN EFNLPPIVAK EIVASCDKCQ
00BW1686_8  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPVVAK EIVASCDKCQ
00BW1759_3  GIRKVLFLDG IDKAQGGSEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW1773_2  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPVVAK EIVASCDKCQ
00BW1783_5  GIRKVLFLDG IDKAQEDHEK YHSNWRAMAN EFNLPPIVAK EIVASCDKCQ
00BW1795_6  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAN EFNLPPIVAK EIVASCDKCQ
00BW1811_3  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPVVAK EIVASCDECQ
00BW1859_5  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAN EFNLPPIVAK EIVASCDKCQ
00BW1880_2  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPVVAK EIVASCDKCQ
00BW1921_1  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPVVAK EIVASCDKCQ
00BW2036_1  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW2063_6  GIRKVLFLDG IDKAQEDHER YHSNWRTMAS EFNLPPIVAK EIVASCDKCQ
00BW2087_2  GVRRVLFLDG IDKAQEDHEK YHSNWRAMAN EFNLPPIVAK EIVASCDKCQ
00BW2127_2  GIRKVLFLNG IDKAQEEHEK YHSNWRAMAN EFNLPPVVAK EIVASCDKCQ
00BW2128_3  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW2276_7  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPIVAK EIVASCDKCQ
00BW3819_3  GIRKVLFLDG IDKAQEDHEK YHNNWRAMVS EFNLPPIVAK EIVASCDKCQ
00BW3842_8  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW3871_3  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW3876_9  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW3886_8  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW3891_6  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPVVAK EIVASCDKCQ
00BW3970_2  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
00BW5031_1  GIRKVLFLDG IDKAQEEHEK YHSNWRSMAS DFNLPPIVAK EIVASCDKCQ
 96BW01B21  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS EFNLPPIVAK EIVASCDKCQ
  96BW0407  GIRKVLFLDG IDKAQEEHEK YHCNWRAMAS DFNLPPVVAK EIVASCDKCQ
  96BW0502  GIRKMLFLDG IDKAQEEHEK YHNNWRAMAD EFNLPPIVAK EIVASCDKCQ
 96BW06_J4  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
 96BW11_06  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
  96BW1210  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPIVAK EIVASCDKCQ
 96BW15B03  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
```

```
 96BW16_26  GIRKVLFLDG IDKAQEDHEK YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
 96BW17A09  GIRKVLFLDG IDKAQEDHEK YHGNWRAMAN EFNLPPIVAK EIVASCDKCQ
 96BWMO1_5  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
 96BWMO3_2  GIRKILFLDG IDKAQEEHEK YHNNWRAMAS EFNLPPIVAK EIVASCDKCQ
98BWMC12_2  GIRKVLFLDG IDKAQDEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
98BWMC13_4  GIRKVLFLDG IDKAQDDHER YHSNWRAMAN EFNLPPIVAK EIVASCDKCQ
98BWMC14_a  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
98BWMO14_1  GIRKVLFLDG IDKAQEEHEK YHSNWRARAS EFNLPPIVAK EIVASCDKCQ
98BWMO18_d  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
98BWMO36_a  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS EFNLPPIVAK EIVASCDKCQ
98BWMO37_d  GIRKVLFLDG IDKAQEDHEK YHSNWRAMAN EFNLPPIVAK EIVASCDKCQ
99BW3932_1  GIRKVLFLDG IDKAREEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
99BW4642_4  GIRKVLFLDG IDKAQDEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
99BW4745_8  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPVVAK EIVASCDKCQ
99BW4754_7  GIRKILFLDE IDKAQDEHEK YHSNWRAMAS EFNLPPVVAK EIVACCDKCQ
99BWMC16_8  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
A2_CD_97CD  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAH DFNLPPIVAK EIVASCDKGQ
A2_CY_94CY  GIRKVLFLDG IDKAQEEHER YHSNWRAMAH DFNLPPVVAK EIVASCDKCQ
A2D___97KR  GIRKVLFLDG IDKAQDEHEK YRSNWRAMAH DFNLPPVVAK EIVACCDKCQ
A2G_CD_97C  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPIVAK EIVASCDKGQ
A_BY_97BL0  XIRXVLFLEX IDKAQEXHEK YHSN.KAMAS DFNLPPIVAX EIVASXDKCQ
A_KE_Q23_A  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVASCDKCQ
A_SE_SE659  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPVIAK EIVASCDKCQ
A_SE_SE725  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPVIAK EIVASCDKCQ
A_SE_SE753  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVASCDKCQ
A_SE_SE853  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPVIAK EIVASCNKCQ
A_SE_SE889  GVRRILFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVASCDKCQ
A_SE_UGSE8  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVASCDKCQ
A_UG_92UG0  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
A_UG_U455_  GIRKVLFLDG IDKAQEDHEK YHCNWRAMAS DFNLPPVVAK EIVASCNKCQ
AC_IN_2130  GIRRVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPIIAK EIVASCNQCQ
AC_RW_92RW  GIRRVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
AC_SE_SE94  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
ACD_SE_SE8  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPVIAK EIVASCDKCQ
ACG_BE_VI1  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPVIAK EIVASCDKCQ
AD_SE_SE69  GIRKILFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPVVAK EIVASCDKCQ
AD_SE_SE71  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
ADHK_NO_97  GIRKVLFLDG IDKAQEAHEK YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
ADK_CD_MAL  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
AG_BE_VI11  GIRKVLFLDG IDKAQADHER YHXNWGAMAS DFNLPPIVAK EIVASCDKCQ
AG_NG_92NG  GIRKVLFLDG IDKAQEDHER YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
AGHU_GA_VI  GIRKVLFLDG IDKAQEDHER YHSNWKAMAS DFNLPPIVAK EIVASCDKCQ
AGU_CD_Z32  GIRKVLFLDG IDKAQEEHER YHCNWRAMAS DFNLPPIVAK EIVASCDKCQ
AJ_BW_BW21  GIRKVLFLDG IDKAQEEHEK YHSKWRAMAS DFNLPPIVAK EIVASCDKCQ
B_AU_VH_AF  GIRKVLFLDG IDKAQDDHEI YHSNWRAMAN DFYLPPIVAK DIVASCDKCQ
B_CN_RL42_  GIRKVLFLDG IDKAQEEHEK YHCNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_DE_D31_U  GIRKVLFLDG IDKAQDEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_DE_HAN_U  GIRKVLFLDG IDKAQEDHEK YHSNWRAMAN DFNLPPVVAK EIVASCDKCQ
B_FR_HXB2_  GIRKVLFLDG IDKAQDEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_GA_OYI__  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_GB_CAM1_  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_GB_GB8_A  GIRKILFLDG IDKAQEDHEK YHSNWRAMAN DFNLPPVVAK EIVASCDKCQ
B_GB_MANC_  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_KR_WK_AF  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAG DFNLPPVVAK EIVACCDKCQ
B_NL_3202A  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_TW_TWCYS  GIRKVLFLDG IDRAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_US_BC_L0  GVRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_US_DH123  GIRRVLFLDG IEKAQEEHEK YHSNWRAMAS EFNLPAVVAK EIVACCDKCQ
B_US_JRCSF  GIRKVLFLDG IDKAQEDHEK YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
B_US_MNCG_  GIRKVLFLDG IDKAQEDHEK YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
B_US_P896_  GIRKVLFLDG IDKAQEEHEK YHTNWRAMAS DFNLPPVVAK EIVASCNKCQ
```

```
B_US_RF_M1  GIRKVLFLDG IDKAQDEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_US_SF2_K  GIRKVLFLNG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_US_WEAU1  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
B_US_WR27_  GIRKVLFXDG IDXAQEDHEK YHSNWRAMAG EFNLPPVXAK EIVASCDKCQ
B_US_YU2_M  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVACCDKCQ
BF1_BR_93B  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNIPPVVAK EIVASCDKCQ
C_BR_92BR0  GIRKVLFLDG INKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
C_BW_96BW0  GIRKVLFLDG IDKAQEEHEK YHCNWRAMAS EFNLPPIVAK EIVASCDKCQ
C_BW_96BW1  GIREVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
C_BW_96BW1  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPIVAK EIVASCDKCQ
C_BW_96BW1  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
C_ET_ETH22  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAN EFNIPPVVPK EIVACCDKCQ
C_IN_93IN1  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDQCQ
C_IN_93IN9  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDQCQ
C_IN_93IN9  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPIVAK EIVASCGQCQ
C_IN_94IN1  GIRRVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDQCQ
C_IN_95IN2  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPVVAK EIVASCDQCQ
CRF01_AE_C  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVANCDKCQ
CRF01_AE_C  GIRKVLFLDG IDKAQEDHER YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
CRF01_AE_C  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVANCDKCQ
CRF01_AE_T  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVANCDKCQ
CRF01_AE_T  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVANCDKCQ
CRF01_AE_T  GIRKVLFLDG INKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVANCDKCQ
CRF01_AE_T  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVANCDKCQ
CRF01_AE_T  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPVVAK EIVANCDKCQ
CRF01_AE_T  GIRKVLFLDG IDKAQEEHER YHSNWRTMAS DFNLPPIVAK EIVTNCDKCQ
CRF02_AG_F  GIRKVLFLDG IDKAQEEHGR YHSNWRAMAS DFNLPPIVAK EIVANCDKCQ
CRF02_AG_F  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPIIAK EIVACCDQCQ
CRF02_AG_G  GFRKILFLDG LDKAQEEHEK FHSNWRAMAS DFNLPPIVAK EIVACCDKCQ
CRF02_AG_N  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
CRF02_AG_S  GIRKVLFLDG IDKAQEEHER YHSNWKAMAS DFNLPPIVAK EIVASCDKCQ
CRF02_AG_S  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
CRF03_AB_R  GIREVLFLDG IDKAQEEHEK YHSNWRAMAS EFNLPPIVAK EIVASCDKCQ
CRF03_AB_R  GIRKVLFLDG IDKAQEAHEK YHGNWRAMAS DFNLPPVVAK EIVASCDKCQ
CRF04_cpx_  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
CRF04_cpx_  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPSVVAK EIVASCNKCQ
CRF04_cpx_  GIRKVLFLDG IDKAQEDHEK YHSNWRAMAS DFNLPPVVAK EIVASCNKCQ
CRF05_DF_B  GIRKVLFLDG IDKAQEEHEK YHNNWKAMAS DFNLPPVVAK EIVASCNKCQ
CRF05_DF_B  GIRKILFLDG IDKAQDEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
CRF06_cpx_  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPVVAK EIVASCDKCQ
CRF06_cpx_  GIRKVLFLDG IDKAQEDHER YHSNWRAMAN DFNLPPIVAK EIVASCDKCQ
CRF06_cpx_  GIRKVLFLDG IDKAQEDHER YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
CRF06_cpx_  GIRKVLFLDG IDKAQEDHER YHSNWRAMAS DFNLPPILAK EIVACCDKCQ
CRF11_cpx_  GIRKVLFLDG IDKAQEEHER YHSNWRAMAN DFNLPPIVAK EIVASCDKCQ
CRF11_cpx_  GIRKVLFLDG IDKAQEDHDR YHSNWRTMAS DFNLPPIVAK EIVASCDKCQ
D_CD_84ZR0  GIRKVLFLDG IDKAPEGHER YHSNWRAMAS DFNLPPVIAK EIVANCDKCQ
D_CD_ELI_K  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
D_CD_NDK_M  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPVVAK EIVASCDKCQ
D_UG_94UG1  GIRKILFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPVVAK EIVASCDKCQ
F1_BE_VI85  GVRKILFLDG IDKAQEEHEK YHNNWRAMAS EFNLPPVVAK EIVASCDKCQ
F1_BR_93BR  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPIVAK EIVASCDKCQ
F1_FI_FIN9  GIRKILFLDG IDKAQEEHEK YHNNWRAMAS DFNIPAVVAK EIVASCDKCQ
F1_FR_MP41  GIRKILFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPVVAK EIVASCDKCQ
F2_CM_MP25  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPVVAK EIVASCDKCQ
F2KU_BE_VI  GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
G_BE_DRCBL  GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPIVAK EIVASCDKCQ
G_NG_92NG0  GIRKVLFLDG IDKAQEEHER YHSNWKAMAS DFNLPPIVAK EIVASCDKCQ
G_SE_SE616  GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
H_BE_VI991  GIRKVLFLDG IDKAQEEHER YHNNWRAMAS DFNLPPIVAK EIVASCDKCQ
H_BE_VI997  GIRKVLFLDG IDKAQVQHEK YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ
H_CF_90CF0  GVRKVLFLDG IDKAQEAHER YHNNWRAMAS EFNLPPIVAK EIVASCDKCQ
            GVRKVLFLDG IDKAQEEHER YHNNWRAVAS DFNLPPIVAK EIVASCDKCQ
```

```
J_SE_SE702   GIRKVLFLDG IDKAQEEHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
J_SE_SE788   GIRKVLFLDG IDKAQEDHEK YHSNWRAMAS DFNLPPVVAK EIVASCDKCQ
K_CD_EQTB1   GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPVVAK EIVASCDKCQ
K_CM_MP535   GIRKVLFLDG IDKAQEEHEK YHNNWRAMAS DFNLPPVVAK EIVASCDKCQ
N_CM_YBF30   GIRKILFLDG IEKAQEDHDR YHSNWKAMAS DFNLPPIVAK EIVASCDKCQ
O_CM_ANT70   DIRRVLFLEG IDQAQEDHEK YHSNWKALAS EFGLPPVVAK EIIASCPKCH
O_CM_MVP51   DIRRVLFLEG IDQAQEDHEK YHSNWRALAS DFGLPPVVAK EIIASCPKCH
O_SN_99SE_   DIRRVLFLEG IDQAQEDHEK YHSNWRALAS DFGLPPIVAK EIIASCPKCH
O_SN_99SE_   DIRRVLFLEG IDQAQEDHEK YHSNWRALAS DFGLPPVVAK EIIANCPQCH
U_CD___83C   GIRKVLFLDG IDKAQEEHER YHSNWRAMAS DFNLPPIVAK EIVASCDKCQ

                    801                                           850
00BW0762_1   LKGEAIHGQV DCSPGIWQLD CTHLEGKTIL VAVHVASGYM EAEVIPAETG
00BW0768_2   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYM EAEVIPAETG
00BW0874_2   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
00BW1471_2   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
00BW1616_2   LKGEATHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW1686_8   LKGEAIHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
00BW1759_3   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPTETG
00BW1773_2   LKGEAIHGQV DCSPGIWQLD CTHLEGKVII VAVHVASGYI EAEVIPAETG
00BW1783_5   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW1795_6   LKGEAIHGQV DCSPGIWQLD CTHLEGKIVL VAVHVASGYL EAEVIPTETG
00BW1811_3   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW1859_5   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYM EAEVIPAETG
00BW1880_2   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVTPEETG
00BW1921_1   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW2036_1   LKGEAIHGQV DCSPGIWQLD CTHLEGKIII VAAHVASGYI EAEVIPAETG
00BW2063_6   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYV EAEVIPAETG
00BW2087_2   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW2127_2   LKGEATHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW2128_3   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW2276_7   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW3819_3   LKGEATHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYM EAEVIPAETG
00BW3842_8   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAIHVASGYM EAEVTPAETG
00BW3871_3   LKGEATHRQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW3876_9   QKGEAIHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
00BW3886_8   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
00BW3891_6   LKGEAIHGQV DCSPGIWQLD CTHLEGKTIL VAVHVASGYI EAEVIPAETG
00BW3970_2   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
00BW5031_1   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
 96BW01B21   LKGEAMHGQV DCSPGVWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
  96BW0407   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYM EAEVIPAETG
  96BW0502   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
 96BW06_J4   SKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYM EAEVIPAETG
 96BW11_06   LKGEAIHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
  96BW1210   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
 96BW15B03   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVRVASGYI EAEVIPAETG
 96BW16_26   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
 96BW17A09   LKGEAMHGQV DCSPGMWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
 96BWMO1_5   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYM EAEVIPAETG
 96BWMO3_2   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYM EAEVIPAETG
98BWMC12_2   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYM EAEVITAETG
98BWMC13_4   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
98BWMC14_a   LKGEAIHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVISAETG
98BWMO14_1   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
98BWMO18_d   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
98BWMO36_a   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
98BWMO37_d   LKGEAIHGQV DCSPGIWQLD CTHLEGKIII VAVHVASGYV EAEVIPAETG
99BW3932_1   LKGEATHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
99BW4642_4   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
99BW4745_8   LKGEAIHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
```

```
99BW4754_7  QKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
99BWMC16_8  LKGEAIHGQV DCSPGVWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
A2_CD_97CD  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIV VAVHVASGYI EAEVIPAETG
A2_CY_94CY  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPTETG
A2D___97KR  VKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
A2G_CD_97C  LKGEAMHGQV DCSPGVWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
A_BY_97BL0  LKXEXMHXQV DCSPXIWQLD CTHLXXKVII XAVXVASGYI EAEVIPAETG
A_KE_Q23_A  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIV VAVHVASGYI EAEVIPAETG
A_SE_SE659  LKGEAIHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
A_SE_SE725  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
A_SE_SE753  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
A_SE_SE853  LKGEAMHGQV DCSPGMWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
A_SE_SE889  LKGEAMHGQV DCSPGMWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
A_SE_UGSE8  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
A_UG_92UG0  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
A_UG_U455_  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYV EAEVIPAETG
AC_IN_2130  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
AC_RW_92RW  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
AC_SE_SE94  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHAASGYI EAEVIPAETG
ACD_SE_SE8  IKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
ACG_BE_VI1  LKGEAMHGQV DCSPGIWQID CTHLEGKVII VAVHVASGYM EAEVIPAETG
AD_SE_SE69  LKGEALHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
AD_SE_SE71  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
ADHK_NO_97  LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
ADK_CD_MAL  LKGEAMHGQV DCSPGIWQLD CTHLEGKIII VAVHVASGYI EAEVIPAETG
AG_BE_VI11  LKGEAIHGQV DCSPGIWQLD CTHLEGKIII VAVHVASGYI EAEVIPAETG
AG_NG_92NG  LKGEAMHGQV DCSPGIWQLD CTHLEGKVII VAVHVASGYI EAEVIPAETG
AGHU_GA_VI  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
AGU_CD_Z32  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
AJ_BW_BW21  LKGEAMHGQV DCSPGIWQLD CTHLEGKVII VAVHVASGYM EAEVIPAETG
B_AU_VH_AF  QKGEAMHGQV DCSPGLWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
B_CN_RL42_  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_DE_D31_U  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
B_DE_HAN_U  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPVETG
B_FR_HXB2_  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
B_GA_OYI__  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_GB_CAM1_  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
B_GB_GB8_A  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_GB_MANC_  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
B_KR_WK_AF  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
B_NL_3202A  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_TW_TWCYS  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_US_BC_L0  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_US_DH123  VKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
B_US_JRCSF  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_US_MNCG_  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
B_US_P896_  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
B_US_RF_M1  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_US_SF2_K  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_US_WEAU1  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_US_WR27_  LKGEAMHGQV DCSPGIWXLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
B_US_YU2_M  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
BF1_BR_93B  LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVAGGYI EAEVIPAETG
C_BR_92BR0  LKGEATHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
C_BW_96BW0  LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
C_BW_96BW1  LKGEAIHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
C_BW_96BW1  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVRVASGYI EAEVIPAETG
C_BW_96BW1  LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
C_ET_ETH22  LKGEAIHGQV NCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
C_IN_93IN1  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
C_IN_93IN9  LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
```

```
C_IN_93IN9   QKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
C_IN_94IN1   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYM EAEVIPAETG
C_IN_95IN2   LKGEATHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYM EAEVIPAETG
CRF01_AE_C   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF01_AE_C   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF01_AE_C   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF01_AE_T   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF01_AE_T   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF01_AE_T   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF01_AE_T   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF01_AE_T   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF01_AE_T   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF02_AG_F   LKGEAMHGQV DCGPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF02_AG_F   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF02_AG_G   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF02_AG_N   MKGEAMHGQV DCGPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF02_AG_S   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF02_AG_S   LKGEAMHGQV DCSPGIWQID CTHLEGKIII VAVHVASGYI EAEVIPAETG
CRF03_AB_R   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF03_AB_R   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF04_cpx_   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIM VAVHVASGYI EAEVIPAETG
CRF04_cpx_   LKGEAMHGQX DCSPGIWQLD CTHLEGKIIL VPVHVASGYI EAEVIPAETG
CRF04_cpx_   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF05_DF_B   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF05_DF_B   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
CRF06_cpx_   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF06_cpx_   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF06_cpx_   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF06_cpx_   LKGEPMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
CRF11_cpx_   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIM VAVHVASGYI EAEVIPTETG
CRF11_cpx_   LKGEAMHGQV DCSPRIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
D_CD_84ZR0   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
D_CD_ELI_K   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
D_CD_NDK_M   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
D_UG_94UG1   VKGEALHGQV DCSPGIWQLD CTHLEGKGIL VAVHVASGYI EAEVIPAETG
F1_BE_VI85   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASEYI EAEVIPAETG
F1_BR_93BR   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYL EAEVIPAETG
F1_FI_FIN9   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
F1_FR_MP41   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
F2_CM_MP25   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
F2KU_BE_VI   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
G_BE_DRCBL   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
G_NG_92NG0   LKGEAMHGQV DCSPGIWQLD CTHLEGKIII VAVHVASGYI EAEVIPAETG
G_SE_SE616   LKGEAMHGQV DCSPGIWQLD CTHLEGKIII VAVHVASGYI EAEVIPAETG
H_BE_VI991   LKGEAMHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
H_BE_VI997   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI XPEVIPAETG
H_CF_90CF0   LKGEAMHGQV DCSPGIWQLD CTHLEGQVIL VAVHVASGYI EAEVIPAETG
J_SE_SE702   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
J_SE_SE788   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYI EAEVIPAETG
K_CD_EQTB1   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
K_CM_MP535   LKGEAIHGQV DCSPGIWQLD CTHLEGKIIL VAVHVASGYI EAEVIPAETG
N_CM_YBF30   LKGEAMHGQV NCSPGVWQLD CTHLEGKIIL VAVHVASGYL EAEVIPAETG
O_CM_ANT70   IKGEAIHGQV DCSPEVWQID CTHMEGKIII VAVHVASGFI EAEVIPAETG
O_CM_MVP51   IKGEATHGQV DYSPEIWQMD CTHLEGKIII VAVHVASDFI EAEVIPAETG
O_SN_99SE_   IKGEAIHGQV DCSPEVWQMD CTHLEGKIII VAVHVASGFI EAEVIPAETG
O_SN_99SE_   IKGEAIHGQV DYSPEVWQID CTHLEGKIII VAVHVASGFI EAEVIPAETG
U_CD___83C   LKGEAMHGQV DCSPGIWQLD CTHLEGKVIL VAVHVASGYL EAEVIPAETG

             851                                                 900
00BW0762_1   QETAYYILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW0768_2   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
```

```
00BW0874_2   QDTAYYILKL AGRWPVKVIH TDNGSNFTST AVKAACWWAD IQQEFGIPYN
00BW1471_2   QEAAYFILKL AGRWPVQIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW1616_2   QETAYYILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW1686_8   QETAYYILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW1759_3   QETAYLILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW1773_2   QETAYYILKL AGRWPVKVIH TDNGSNFTST TVKAACWWAG IQQEFGIPYN
00BW1783_5   QETAYFILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG VQQEFGIPYN
00BW1795_6   QETAYYILKL AGRWPVKVIH TDNGSNFTSA TVKAACWWAG IQQEFGIPYN
00BW1811_3   QETAYFLLKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW1859_5   QETAYYILKL AGRWPVKVIH TDNGSNFTSN AMKAACWWAG IQQEFGIPYN
00BW1880_2   QETAYYILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW1921_1   QETAYYILKL AGRWPVKVIH TDNGRNFTSA AVKAACWWAG IQQEFGIPYN
00BW2036_1   QETAYFILKL AGRWPVKVIH TDNGSNFTST AVKAACWWAG IQQEFGIPYN
00BW2063_6   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW2087_2   QETAYYILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW2127_2   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW2128_3   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW2276_7   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYS
00BW3819_3   QETAYFILKL AGRWPVKVIH TDNGSNFTST TVKAACWWAG IQQEFGIPYN
00BW3842_8   QETAYYILKL AGRWPVKVIH TDNGSNFTSN AVKAACWWAG IQQEFGIPYN
00BW3871_3   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW3876_9   QETAYYILKL AGRWPVKVIH TDNGSNFTST VVKAACWWAG IQQEFGIPYN
00BW3886_8   QETAYYILKL AGRWPVKVIH TDNGSNFTSS AVKAACWWAG TQQGFGIPYN
00BW3891_6   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
00BW3970_2   QETAYYILKL AGRWPVKVIH TDNGSNFTSN AVKAACWWAG IQQEFGIPYN
00BW5031_1   QETVYFILKL AGRWPVKAIH TDNGSNFTST AVKAACWWAG IKQEFGIPYN
 96BW01B21   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG INQEFGIPYN
  96BW0407   QETAYYILKL AGRWPVKVIH TDNGTNFTSA AVKAACWWAG IQQEFGIPYN
  96BW0502   QETAYFILKL AGRWPVKVIH TDNGSNFTST AVKAACWWAG IQQEFGIPYN
 96BW06_J4   QDTAYYILKL AGRWPVRVIH TDNGTNFTSA AVKAACWWAG IQQEFGIPYN
 96BW11_06   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
  96BW1210   QETAYFILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
 96BW15B03   QETAYYILKL AGRWPVKVIH TDNGSNFTSN AVKAACWWAG TQQEFGIPYN
 96BW16_26   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
 96BW17A09   QETADFILKL AGRWPVQIIH TDNGSNFTSA TVKAACWWAG IQQEFGIPYN
 96BWMO1_5   QETAYFLLKL AGRWPVKIIH TDNGSNFTST AVKAACWWAG IQQEFGIPYN
 96BWMO3_2   QDTAYFLLKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
98BWMC12_2   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AMKAACWWAG IQQEFGIPYN
98BWMC13_4   QETAYFILKL AGRWPVKVIH TDNGSNFTSG AVKAACWWAG IQQEFGIPYN
98BWMC14_a   QETAYYILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
98BWMO14_1   QETAYYILKL AARWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
98BWMO18_d   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
98BWMO36_a   QETAYYILKL AGRWSVKTIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
98BWMO37_d   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
99BW3932_1   QETAYYILKL AGRWPVKVIH TDNGSNFTST AVKAACWWAG IQQEFGIPYN
99BW4642_4   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
99BW4745_8   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
99BW4754_7   QETAYYILKL AGRWPVKVIH TDNGSNFTST AVKAACWWAG IQQEFGIPYN
99BWMC16_8   QETAYYILKL AGRWPVKTIH TDNGSNFTST AVKAACWWAG IQQEFGIPYN
A2_CD_97CD   QDTAYYMLKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
A2_CY_94CY   QETAYFLLKL AGRWPVKVIH TDNGPNFTSA AVKAACWWAD VQQEFGIPYN
A2D___97KR   QDTAYFILKL AGRWPVKVIH TDNGPNFISA TVKAACWWAG IQQEFGIPYN
A2G_CD_97C   QETAYFILKL AGRWPVKVIH TDNGPNFISA PVKAACWWAG IQQEFGIPYN
A_BY_97BL0   QETAYFILKL AGGWPVKVIH TDNGSNFTSA AVKAACWWAN VQQEFGIPYN
A_KE_Q23_A   QETAYFLLKL AGRWPVKVVH TDNGPNFTSS AVKAACWWAN VTQEFGIPYN
A_SE_SE659   QETAYFLLKL AGRWPVKIVH TDNGSNFTSA AVKAACWWAN. IQQEFXIPYN
A_SE_SE725   QETAYFLLKL AGRWPVKVVH TDNGSNFTSA AVKAACWWAN IQQEFGIPYN
A_SE_SE753   QETAYFLLKL AGRWPVKIVH TDNGSNFTSA AVKAACWWAN IQQEFGIPYN
A_SE_SE853   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AFKAACWWAS IQQEFGIPYN
A_SE_SE889   QETAYFLLKL AGRWPVKVVH TDNGSNFTSA AVKAACWWAN VKQEFGIPYN
A_SE_UGSE8   QETAYFLLKL AGRWPVKVVH TDNGSNFTSA AFKAACWWAN VQQEYGIPYN
             QEAAYFLLKL AGRWPVKVVH TDNGSNFTSA AFKAACWWAN VQQEFGIPYN
```

```
A_UG_92UG0   QETAYFLLKL AGRWPVKVVH TDNGSNFTSA AVKAACWWAN VKQEFGIPYN
A_UG_U455_   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAVCWWAN IQQEFGIPYN
AC_IN_2130   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
AC_RW_92RW   QETAYFILKL AGRWPVKVIH TDNGSNFTSN TVKAACWWAG IQQEFGIPYN
AC_SE_SE94   QETAYFLLKL AGRWPVRRVH TDNGSNFTSA AVKAACWWAN IQQEFGIPYN
ACD_SE_SE8   QETAYFLLKL AGRWPVKVVH TDNGSNFTSA AFKAACWWAS VQQEFGIPYN
ACG_BE_VI1   QETAYFLLKL AGRWPVRVIH TDNGSNFTSA AVKAACWWAN VTQEFGIPYN
AD_SE_SE69   QETAYFLLKL AGRWPVRVVH TDNGSNFTST AVKAACWWAG IKQEFGIPYN
AD_SE_SE71   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
ADHK_NO_97   QETAYFILKL AGRWPVKVIH TDNGSNFISA AVKAACWWAD IKQEFGIPYN
ADK_CD_MAL   QETAYFILKL AGRWPVKVVH TDNGSNFTSA AVKAACWWAN IKQEFGIPYN
AG_BE_VI11   QETAYFILKL AGRWPVKILH TDNGSNFISA AVKAACWWAD IKQEFGIPYN
AG_NG_92NG   QETAYFLLKL AGRWPVKVIH TDNGSNFTSA AMKAACWWAN IQQEFGIPYN
AGHU_GA_VI   QETAYFILKL AGRWPVKVIH TDNGTNFTSA AVKAACWWAN VTQEFGIPYN
AGU_CD_Z32   QETAYFILKL AGRWPVKVVH TDNGSNFTSA AVKAACWWAN ITQEFGIPYN
AJ_BW_BW21   QETAYFLLKL AGRWPVTVIH TDNGSNFTSA AVKAACWWAG VRQEFGIPYN
B_AU_VH_AF   QETAYFLLKL AGRWPVKTVH TDNGPIFIST AVKAACWWAG IKQEFGIPYN
B_CN_RL42_   QETAYFLLKL AGRWPVKTIH TDNGRNFTSN SVKAACWWAG IKQEFGIPYN
B_DE_D31_U   QETAYFILKL AGRWPVKTIH TDNGSNFTST TVKAACWWAG VKQEFGIPYN
B_DE_HAN_U   QETAYFLLKL AGRWPVKTVH TDNGPNFTST TVKAACWWAG IKQEFGIPYN
B_FR_HXB2_   QETAYFLLKL AGRWPVKTIH TDNGSNFTGA TVRAACWWAG IKQEFGIPYN
B_GA_OYI__   QETAYFILKL AGRWPVKTIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
B_GB_CAM1_   QETAYFLLKL AGRWPVKTIH TDNGGNFIST TVKAACWWAG IKQEFGIPYN
B_GB_GB8_A   QDTAYFILKL AGRWPVKTIH TDNGRNFTST TVKAACWWAG IKQEFGIPYN
B_GB_MANC_   QETAYFLLKL AGRWPVTTIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
B_KR_WK_AF   QETAYFLLKL AGRWPVKTIH TDNGSNFTSN TVKAACWAR. IKQEFGIPYN
B_NL_3202A   QETAYFILKL AGRWPVTTIH TDNGSNFTSA TVKAACWWAG IKQEFGIPYN
B_TW_TWCYS   QETAYFLLKL AGRWPVKTIH TDNGSNFTSA AVKAACWWAG IKQKFGIPYN
B_US_BC_L0   QETAYFLLKL AGRWPVKTIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
B_US_DH123   QETAYFILKL AGRWPVKTIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
B_US_JRCSF   QETAYFLLKL AGRWPVTTIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
B_US_MNCG_   QETAYFLLKL AGRWPVKTIH TDNGPNFTST TVKAACWWTG IKQEFGIPYN
B_US_P896_   QETAYFLLKL AGRWPVKTIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
B_US_RF_M1   QETAYFILKL AGRWPVKVIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
B_US_SF2_K   QETAYFLLKL AGRWPVKTIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
B_US_WEAU1   QETAYFILKL AGRWPVKTIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
B_US_WR27_   QETAYFILKL AXRWPVXTIH TDNGSNFIST TVXAAXWWAG IXQEFGIPYN
B_US_YU2_M   QETAYFLLKL AGRWPVTTIH TDNGSNFTSA TVKAACWWAG IKQEFGIPYN
BF1_BR_93B   QETAYFLLKL AGRWPVKTIH TDNGSNFTST TVKAACWWAG IKQEFGIPYN
C_BR_92BR0   QETAYFILKL AGRWPVKVIH TDNGSNFISN TVKAACWWAG IQQEFGIPYN
C_BW_96BW0   QETAYYILKL AGRWPVKVIH TDNGSNFTST AVKAACWWAG IQQEFGIPYN
C_BW_96BW1   QETAYYILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
C_BW_96BW1   QETAYFILKL AGRWPVKVIH TDNGSNFTSN AVKAACWWAG TQQEFGIPYN
C_BW_96BW1   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
C_ET_ETH22   QETAYFLLKL AGRWPVRVIH TDNGSNFTSN AVKAACWWAG IQQEFGIPYN
C_IN_93IN1   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
C_IN_93IN9   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAS IQQEFGIPYN
C_IN_93IN9   QETAYYILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
C_IN_94IN1   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
C_IN_95IN2   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
CRF01_AE_C   QETAYFLLKL AG.RPVKVIH TDNGSNFTSA AMKAACWWAN VQQEFGIPYN
CRF01_AE_C   QETAYFLLKL AGRWPVRVIH TDNGSNFTSA AVKAACWWAN VQQEFGIPYN
CRF01_AE_C   QETAYFLLKL AGRWPVKVIH TDNGSNFTSA TVKAACWWAN VQQEFGIPYN
CRF01_AE_T   QETAYFLLKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VRQEFGIPYN
CRF01_AE_T   QETAYFLLKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VRQEFGIPYN
CRF01_AE_T   QETAYFLLKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VRQEFGIPYN
CRF01_AE_T   QETAYFLLKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VRQEFGIPYN
CRF01_AE_T   QETAYFLLKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VQQEFGIPYN
CRF01_AE_T   QETAYFLLKL AGRWPVKVVH TDNGSNFTSA AVKAACWWAN VRQEFGIPYN
CRF02_AG_F   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VTQEFGIPYN
CRF02_AG_F   QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VTQEFGIPYN
```

```
CRF02_AG_G  QETAYFILKL AGRWPVKIIH TDNGSNFTSA AVRAACWWAN VTQEFGIPYN
CRF02_AG_N  QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VTQEFGIPYN
CRF02_AG_S  QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VTQEFGIPYN
CRF02_AG_S  QETAYFILKL AGRWP.KVVH TDNGSNFTSA AVKAACWWPK LHQELGIYN.
CRF03_AB_R  QETAYFVLKL AGRWPVKIIH TDNGSNFTST AVKAACWWAG IKQEFGIPYN
CRF03_AB_R  QETAYFVLKL AGRWPVKVIH TDNGSNFIST AVKAACWWAG IKQEFGIPYN
CRF04_cpx_  QETAYFILKL AGRWPVKMIH ADNGPNFTSA AVKAACWWAG IKQEFGIPYN
CRF04_cpx_  QETAYFILKL AGRWPVKIIH TDNGPNFTSA AVKAACWWAD INQEFGIPYN
CRF04_cpx_  QETAYFILKL AGRWPVKIIH TDNGPNFTSA AVKAACWWAD VQQEFGIPYN
CRF05_DF_B  QDTAYFILKL AGRWPVKMIH TDNGPNFTSA AVKAACWWAN IQQEFGVPYN
CRF05_DF_B  QETAYFILKL AGRWPVKMVH TDNGSNFTSG AVKAACWWAG IQQEFGIPYN
CRF06_cpx_  QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAG IKQEFGIPYN
CRF06_cpx_  QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN ITQEFGIPYN
CRF06_cpx_  QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN ITQEFGIPYN
CRF06_cpx_  QETAYFILKL AGRWPVKVIH TDNGSNFTST AVKAACWWAN VTQEFGIPYN
CRF11_cpx_  QETAYFILKL AGRWPVKVIH TDNGSNFTSA AVKAACWWAN VTQEFGIPYN
CRF11_cpx_  QETAYFILKL AGRWPVKVIH TDNGSNFTSS TVKAACWWAN IQQEFGIPYN
D_CD_84ZR0  QEAAYFLLKL AGRWPVKVVH TDNGSNFTSN AVKAACWWAG IQQEFGIPYN
D_CD_ELI_K  QETAYFLLKL AGRWPVKVVH TDNGSNFTSA TVKAACWWAG IKQEFGIPYN
D_CD_NDK_M  QETAYFLLKL AGRWPVKVVH TDNGSNFTSA AVKAACWWAG IKQEFGIPYN
D_UG_94UG1  QETAYFLLKL AGRWPVKVVH TDNGSNFTSA TVKAACWWAG IKQEFGIPYN
F1_BE_VI85  QETAYFILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IKQEFGIPYN
F1_BR_93BR  QETAYFLLKL AGRWPVKTIH TDNGSNFTSA AVKASCWWAG IQQEFGIPYN
F1_FI_FIN9  QDTAYFILKL AGRWPVKMIH TDNGTNFTSA TVKAACWWAG IQQEFGIPYN
F1_FR_MP41  QETAYFILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN
F2_CM_MP25  QEAAYFILKL AGRWPVKIIH TDNGSNFTSS AVKAACWWAG IQQEFGIPYN
F2KU_BE_VI  QETAFFILKL AGRWPVKIIH TDNGSNFTSA VVKAACWWAG IQQEFGIPYN
G_BE_DRCBL  QETAYFILKL AGRWPVKIIH TDNGSNFISA TVKAACWWAG IQQEFGISYN
G_NG_92NG0  QETAYFILKL AGRWPVKIIH TDNGSNFTSA AVKAACWWAS ITQEFGIPYN
G_SE_SE616  QETAYFILKL AGRWPVKVIH TDNGPNFISA AVKAACWWAN ITQEFGIPYN
H_BE_VI991  QETAYFILKL AGRWPVTVIH TDNGSNFTSA AVKAACWWAN ITQEFGIPYN
H_BE_VI997  QETAYFILKL AGRWPVKMIH TDNGSNFTSA AVKAACWWAD IHQEFGIPYN
H_CF_90CF0  KETAYFLLKL ASRWPVKVIH TDNGSNFTST AVKAACWWAD IQQDFGIPYN
J_SE_SE702  QEAAFFILKL AGGWPVKAIH TDNGSNFTSA AVKAACWWAD IQQEFGIPYN
J_SE_SE788  QEAAFFILKL AGRWPVKVIH TDNGSNFTSG AVKAACWWAD IKQEFGIPYN
K_CD_EQTB1  QETAYFILKL AGRWPVKVIH TDNGSNFTSG AVKAACWWAD IKQEFGIPYN
K_CM_MP535  QETAYFILKL AGRWPVRVIH TDNGSNFTSA VVKAACWWAD IKQEFGIPYN
N_CM_YBF30  QETAYFILKL AGRWPVKVIH TDNGTNFTST VVKAACWWAG VKQEFGIPYN
O_CM_ANT70  QETAYFILKL AGRWPVKVIH TDNGSNFTSA TVKAACWWAN IKQEFGIPYN
O_CM_MVP51  QETAYFLLKL AARWPVKVIH TDNGPNFTST TMKAACWWAN IQHEFGIPYN
O_SN_99SE_  QETAYFLLKL AARWPVKVIH TDNGPNFTSA AMKAACWWTG IQHEFGIPYN
O_SN_99SE_  QETAYFLLKL AARWPVKVIH TDNGPNFTSA AMKAACWWAN IKHEFGIPYN
U_CD___83C  QETAYFVLKL AGRWPVTVIH TDNGSNFTSA AVKAACWWAG IQQEFGIPYN

            901                                                  950
00BW0762_1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRRGGIGGY
00BW0768_2  PQSQGVVESM NKELKKIIGQ IRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW0874_2  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRRGGIGGY
00BW1471_2  PQSQGVVKSM NKELKKIIGQ VRDQAEHLKT AVLMAVFIHN FKRKGGIGGY
00BW1616_2  PQSQGVVESM NKELKKIIGQ IRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW1686_8  PQSQGVVESM NKELKKIIGQ VRDQAEHHKT AVQMAVFVHN FKRKGGIGGY
00BW1759_3  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW1773_2  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW1783_5  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW1795_6  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW1811_3  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRRGGIGGY
00BW1859_5  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW1880_2  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW1921_1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW2036_1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW2063_6  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
```

```
00BW2087_2  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW2127_2  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW2128_3  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW2276_7  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW3819_3  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRRGGIGGY
00BW3842_8  PQSQGVVESM NKELKKIIGQ IRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW3871_3  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW3876_9  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW3886_8  PQSQGVVGSM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW3891_6  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW3970_2  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
00BW5031_1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRRGGIGGY
 96BW01B21  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGEIGGY
  96BW0407  PQSQGVVESM NKELEKIIGQ VREQAEHLKT AVQMAVFIHN FKRRGGIGGY
  96BW0502  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGRIGGY
 96BW06_J4  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
 96BW11_06  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
  96BW1210  PQSQGVVESM NKELKKIIGQ VRDQAEYLKT AVQMAVFIHN FKRKGGIGGY
 96BW15B03  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
 96BW16_26  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
 96BW17A09  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
 96BWMO1_5  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVLMAVFIHN FKRKGGIGGY
 96BWMO3_2  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
98BWMC12_2  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
98BWMC13_4  PQSQGVVESM NKELKKIIGQ IRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
98BWMC14_a  PQSQGAVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
98BWMO14_1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
98BWMO18_d  PQSQGVVESM NKELKKIIGQ VRDQAEHHKT AVQMAVFIHN FKRKGGIGGY
98BWMO36_a  PQSQGVVESM NKELKKIIGQ VRDQAEHPKT AVQMAVFIHN FKRKGGIGGY
98BWMO37_d  PQSQGVVESM NKELKKIIGQ VRDRAEHLKT AVQMAVFIHN FKRKGGIGGY
99BW3932_1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
99BW4642_4  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
99BW4745_8  PQSQGVVEPM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
99BW4754_7  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
99BWMC16_8  PQSQGVVEST NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRRGGIGGY
A2_CD_97CD  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
A2_CY_94CY  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
A2D___97KR  PQSQGVVESM NKELKKIIGQ IRDQAEHLKT AVQMAVFIHN SKRKGGIGGY
A2G_CD_97C  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGRY
A_BY_97BL0  XQSQGVVESM NXELKKIIRQ VREQAEYLKT AVQMAVFIHN FKRKGGIGGY
A_KE_Q23_A  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
A_SE_SE659  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
A_SE_SE725  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
A_SE_SE753  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRRGGIGGY
A_SE_SE853  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
A_SE_SE889  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
A_SE_UGSE8  PQSQGVVESM NKELKKIIGQ VREQAEHLRT AVQMAVFIHN FKRKGGIGGY
A_UG_92UG0  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
A_UG_U455_  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
AC_IN_2130  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
AC_RW_92RW  PQSQGVIESM NKELKKIIGQ VRDQAEHLRT AVQMAVFIHN FKRKGGIGGY
AC_SE_SE94  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
ACD_SE_SE8  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
ACG_BE_VI1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
AD_SE_SE69  PQSQGVVESM NKELKKIIGQ VRDQAEHLRT AVQMAVFIHN FKRKGGIGGY
AD_SE_SE71  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
ADHK_NO_97  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AXQMAVFIHN FKRKGGIGGY
ADK_CD_MAL  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
AG_BE_VI11  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
AG_NG_92NG  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
AGHU_GA_VI  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
```

```
AGU_CD_Z32  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
AJ_BW_BW21  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_AU_VH_AF  PQRQGVVDSM NNDLTTIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_CN_RL42_  PQSQGVVESM DKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_DE_D31_U  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_DE_HAN_U  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_FR_HXB2_  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_GA_OYI__  PQSQGVVESM NNELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_GB_CAM1_  PQSQGVVESM NNELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_GB_GB8_A  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_GB_MANC_  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_KR_WK_AF  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_NL_3202A  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_TW_TWCYS  PQSQGVIESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_US_BC_L0  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_US_DH123  PQSQGVVESM NKELKKIIEQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_US_JRCSF  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_US_MNCG_  PQSQGVIESM NKELKKIIGQ VRDQAEHLKR AVQMAVFIHN FKRKGGIGGY
B_US_P896_  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_US_RF_M1  PQSQGVVESM NKQLKQIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_US_SF2_K  PQSQGVVESM NNELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_US_WEAU1  PQSQGVIESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_US_WR27_  PQSQGVVESX NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
B_US_YU2_M  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
BF1_BR_93B  PQSQGVIESM NKELKKIIGQ VRDQAEHLKT AVQTAVFIHN FKRKGGIGGY
C_BR_92BR0  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKR.GGIGGY
C_BW_96BW0  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
C_BW_96BW1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
C_BW_96BW1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
C_BW_96BW1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
C_ET_ETH22  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRRGGIGGY
C_IN_93IN1  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
C_IN_93IN9  PQSQGVVEAM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
C_IN_93IN9  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQVAVFIHN FKRRGGIGGY
C_IN_94IN1  PQSQGVVEAM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
C_IN_95IN2  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF01_AE_C  PQSQGVVESM NKELKKVIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF01_AE_C  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGEY
CRF01_AE_C  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF01_AE_T  PQGQGVVESM NKELKEIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF01_AE_T  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AF.MAVFIHN FKRKGGIG.Y
CRF01_AE_T  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF01_AE_T  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF01_AE_T  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF01_AE_T  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF02_AG_F  PQSQGVVEAM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF02_AG_F  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF02_AG_G  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF02_AG_N  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF02_AG_S  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVLIHN FKRRGGIGGY
CRF02_AG_S  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRRGGIGGY
CRF03_AB_R  PQSQGVVESM NKQLKQTIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF03_AB_R  PQSQGVVESM NKQLKQIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF04_cpx_  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF04_cpx_  PQSQGVVESM NKELKKIIKQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF04_cpx_  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIEGY
CRF05_DF_B  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMTVFIHN FKRKGGIGGY
CRF05_DF_B  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF06_cpx_  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF06_cpx_  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF06_cpx_  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
```

```
CRF06_cpx_  PQSQGVVKSM NKELKKIIGQ IRDQAEHLKT AVQMAVYIHN FKRKGGIGGY
CRF11_cpx_  PQSQGVVESM NKELKKIIGQ IRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
CRF11_cpx_  PQSQGVVESM NKGLKEIIGQ VREQAEHLKT AVQMAVFIHN FKGKGGIGGY
D_CD_84ZR0  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
D_CD_ELI_K  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
D_CD_NDK_M  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRRRGIGGY
D_UG_94UG1  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
F1_BE_VI85  PQSQGVVESI NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
F1_BR_93BR  PQSQGVVESM NKELKKIIGQ IRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
F1_FI_FIN9  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
F1_FR_MP41  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
F2_CM_MP25  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
F2KU_BE_VI  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
G_BE_DRCBL  PQSQGVVESM NKELKKIIGQ IRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
G_NG_92NG0  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
G_SE_SE616  PQSQGVVESM NKELKKIIGQ VGDQAEHLKT AVQMAVFIHN FKRKGGIGGY
H_BE_VI991  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
H_BE_VI997  PQSQGVVESM NKELKKIIGQ VRDQAEHLRT AVQMAVFIHN FKRKGGIGGY
H_CF_90CF0  PQSQGVVESM NKELKKIIGQ VRDQAEHLRT AVQMAVFIHN FKRKGGIGGY
J_SE_SE702  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
J_SE_SE788  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
K_CD_EQTB1  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
K_CM_MP535  PQSQGVVESM NKELKKIIGQ VREQAEHLKT AVQMAVFIHN FKRKGGIGGY
N_CM_YBF30  PQSQGAVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
O_CM_ANT70  PQSQGVVEAM NKELKKIIGQ IRDQAEHLKT AVQMAVFIHN FKRKGGIGGY
O_CM_MVP51  PQSQGVVEAM NKELKSIIQQ VRDQAEHLRT AVQMAVFVHN FKRKGGIGGY
O_SN_99SE_  PQSQGVVEAM NKELKSIIQQ VRDQAEHLKT AVQMAVFVHN FKRKGGIGGY
O_SN_99SE_  PQSQGVVEAM NKELKSIIQQ VRDQAEHLKT AVQMAVFVHN YKRKGGIGGY
U_CD___83C  PQSQGVVESM NKELKKIIGQ VRDQAEHLKT AVQMAVFIHN FKRKGGIGGY

            951                                                1000
00BW0762_1  SAGERIIDII ATDIQTRELQ KRIIQIQNFR VYYRDSRDPI WKGPAKLLWK
00BW0768_2  SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW0874_2  SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW1471_2  SAGERIVDII ASDIQTKELQ KQITKIQNFR VYYRDSREPV WKGPAKLLWK
00BW1616_2  SAGERIVDII ATDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW1686_8  SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW1759_3  SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW1773_2  SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW1783_5  SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYGDSRDPI WKGPAKLLWK
00BW1795_6  SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYRDNRDPI WKGPAKLLWK
00BW1811_3  SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW1859_5  SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW1880_2  SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW1921_1  SAGERIIDII ATDIQTKELQ KQITNIQKFR VYYRDSRDPI WKGPAKLLWK
00BW2036_1  SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW2063_6  SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW2087_2  SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW2127_2  SAGERIIDII ATDIQTKELQ KRIIQIQNFR VYYRDSRDPI WKGPAKLLWK
00BW2128_3  SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW2276_7  SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW3819_3  SAGERIIDII ATDIQTRELQ KQIIQIQNFR VYYRDSRDPI WKGPAKLLWK
00BW3842_8  SAGERIIDII ATDIQTKELQ NQITKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW3871_3  SAGERIIDII ATDIQTKELQ KQITKIQNFR AYYRDSRDPI WKGPAKLLWK
00BW3876_9  SAGERIIDII ATDIQTKELQ NQITKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW3886_8  SAGERIIDII ATDIQTKELQ NQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW3891_6  SAGGRIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW3970_2  SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
00BW5031_1  SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
 96BW01B21  SAGERIIDII ATDIRTKELQ KQIMKIRNFR VYYRDSRDPI WRGPAKLLWK
  96BW0407  SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPV WKGPAKLLWK
```

```
 96BW0502  SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
 96BW06_J4 SAGERIIDII ATDIQTKELQ KQITNIQKFR VYYRDSRDPI WKGPAKLLWK
 96BW11_06 SAGERIIDMI ATDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
  96BW1210 SAGERIIDII ATDIQTTELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
 96BW15B03 SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
 96BW16_26 TAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLIWK
 96BW17A09 SAGERIVDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
 96BWMO1_5 SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYRDSREPV WKGPAKLLWK
 96BWMO3_2 SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDNRDPI WKGPAKLLWK
98BWMC12_2 SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
98BWMC13_4 SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYRDSRDPV WKGPAKLLWK
98BWMC14_a SAGERIIDII ATDIQTQELQ KQIIKIQNFR VYYRDNRDPI WKGPAKLLWK
98BWMO14_1 SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWR
98BWMO18_d SAGERIIDII ATDIQTKELQ KQILNIQKFR VYYRDSRDPI WKGPAKLLWK
98BWMO36_a SAGERIIDII ATDIQTRELQ RQIIKIQIFR VYYRDSRDPI WKGPAKLLWK
98BWMO37_d SAGERIIDII ATDIQTKELQ RQILKIQNFR VYYRDSRDPI WKGPAKLLWK
99BW3932_1 SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSGDPI WKGPAKLLWK
99BW4642_4 SAGERIIDII ATDIQTRELQ KQIIQIQNFR VYYRDSRDPI WKGPAKLLWK
99BW4745_8 SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
99BW4754_7 SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
99BWMC16_8 SAGERIIDII ASDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
A2_CD_97CD SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
A2_CY_94CY SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
A2D___97KR SAGERIIDII ATDIQTKELQ RQITKIQNFR VYYRDSRDPI WKGPAKLLWK
A2G_CD_97C SAGERIIDII ASDIQTKELQ RQITKIQNFR VYYRDSRDPI WKGPAKLLWK
A_BY_97BL0 SAXERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
A_KE_Q23_A SAGERIIDII ATDIQTKELQ KXITKIQNFX VYYRDSRDPI WKGPAKLLXK
A_SE_SE659 SAGERIIDII ATDIQTKELQ KHITKIQNFR VYYRDSRDPL WKGPAKLFWK
A_SE_SE725 SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
A_SE_SE753 SAGERIIDII ATDIQTKELQ KQITKIQKFR VYYRDSRDPI WKGPAKLLWK
A_SE_SE853 SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
A_SE_SE889 SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
A_SE_UGSE8 SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
A_UG_92UG0 SAGERIIDII ASDLQTKELQ KQIT.IQNFR VYYRDSRDPI WKGPAKLLWK
A_UG_U455_ SAGERIIDII ATDIQTKELQ KQITKIQKFR VCYRDSRDPI WKGPAKLLWK
AC_IN_2130 SAGERIIDII ATDIQTKELQ KQISKIQNFR VYYRDSRDPI WKGPAKLLWK
AC_RW_92RW SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
AC_SE_SE94 SPGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
ACD_SE_SE8 SVGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
ACG_BE_VI1 SARERIIDII ASDIQTKELQ KQITKIQKFR VYYRDSRNPI WKGPAKLLWK
AD_SE_SE69 SAGERIIDII ATDIQTKELQ KPITKIQNFR VYYRDSRDPI WKGPAKLLWK
AD_SE_SE71 SAGERIIDII ATDIQTKELQ RQITKIQNFR VYYRDSRDPI WKGPAKLLWK
ADHK_NO_97 SAGERIIDXI ATDIQTKELQ KHITKIQNFR VYYRDSRDPL WKGPAKLLWK
ADK_CD_MAL SAGERIIDMI ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
AG_BE_VI11 SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDNRDPI WKGPAKLLWK
AG_NG_92NG SAGERIIDII ASDIQTKELQ KXITKIXNFR VYYRDSRDPI WKGPAKLLWK
AGHU_GA_VI SAGERIIDII ASDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
AGU_CD_Z32 SAGERIVDII ASDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
AJ_BW_BW21 SAGERIIDMI ASDIQTKELQ NQITKIQNFR VYYRDSRDPI WKGPAKLLWK
B_AU_VH_AF SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
B_CN_RL42_ SAGERIVDII ASDIQTKELQ KQITKVQNFR VYYRDSRDPL WKGPAKLLWK
B_DE_D31_U SAGERIVDII ATDIQTRELQ KQITKIQNFR VYYRGSRDPL WKGPAKLLWK
B_DE_HAN_U SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
B_FR_HXB2_ SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDSREPF WKGPAKLLWK
B_GA_OYI__ SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDSRNPL WKGPAKLLWK
B_GB_CAM1_ SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSREPL WKGPAKLLWK
B_GB_GB8_A SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
B_GB_MANC_ SAGGRIVDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
B_KR_WK_AF SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRNPL WKGPAKLLWK
B_NL_3202A SAGERIVDII ATDIQTKELQ KQVTKIQNFR VYYRDSRDPL WKGPAKLLWK
B_TW_TWCYS SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
```

```
B_US_BC_L0   SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDNKDPL WKGPAKLLWK
B_US_DH123   SAGERIVDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
B_US_JRCSF   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDNRDPI WKGPAKLLWK
B_US_MNCG_   SAGERIVGII ATDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
B_US_P896_   SAGERIVDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
B_US_RF_M1   SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGHAKLLWK
B_US_SF2_K   SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDNKDPL WKGPAKLLWK
B_US_WEAU1   SAGERIIDII ATDIQTQQLQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
B_US_WR27_   SAGERIIDII ATDIQXKXLQ XQXTIXQNXR VYYRDSRDPL WKGPAKLLWK
B_US_YU2_M   SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
BF1_BR_93B   SAGERIVDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPL WKGPAKLLWK
C_BR_92BR0   SAGERIIDII ATDIQTKELQ KQIMKIQNFR VYYRDSRDPI WKGPAKLLWK
C_BW_96BW0   SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPV WKGPAKLLWK
C_BW_96BW1   SAGERIIDMI ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
C_BW_96BW1   SAGERIIDII ATDIQTTELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
C_BW_96BW1   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
C_ET_ETH22   SAGERIIDII ASDIQTKELQ NQILKIQNFR VYYRDSRDPI WKGPAKLIWK
C_IN_93IN1   SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
C_IN_93IN9   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
C_IN_93IN9   SAGERIIDII ATDIQTKELQ KQIIKIHNFR VYYRDSRDPI WKGPAKLLWK
C_IN_94IN1   SAGERIIDII STDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
C_IN_95IN2   SAGERIIDII ATDIQTKELQ KQITKVQNFR VYYRDSRDPI WKGPAKLLWK
CRF01_AE_C   SAGERIIDII ATEIQTKEXQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF01_AE_C   SAGERIIDII ATDIQTKALQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF01_AE_C   SAGERIIDII ATDIQTKELQ KHITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF01_AE_T   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF01_AE_T   SVGERIIDII AADIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF01_AE_T   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF01_AE_T   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF01_AE_T   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF01_AE_T   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF02_AG_F   SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF02_AG_F   SAGERIIDII ASDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF02_AG_G   SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF02_AG_N   SAGERIIDII ASDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF02_AG_S   SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF02_AG_S   SAGERIIDII ASDIQTKELQ KQITNIQKFR VYYRDSRDPI WKGPAKLLWK
CRF03_AB_R   SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF03_AB_R   SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF04_cpx_   SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDSREPI WKGPAKLLWK
CRF04_cpx_   SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF04_cpx_   SAGERIIDII ASDIQTKELQ KQIIKIQNFR VYYRDSKDPI WKGPAKLLWK
CRF05_DF_B   GAGERIIDII TTDIQTKELQ KQIIKIQNFR VYYRDSRDPV WKGPAKLLWK
CRF05_DF_B   SAGEGIIDII STDIQTKELQ KQITKIQNFR VYYRDSRDPV WKGPAKLLWK
CRF06_cpx_   SAGERIIDII ASDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF06_cpx_   SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF06_cpx_   SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF06_cpx_   SAGERIIDII ASDIQTKELQ KQITKIRNFR VYYRDSRDPI WKGPAKLLWK
CRF11_cpx_   SAGERIVDII ATDLQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
CRF11_cpx_   SAGERIIDII ATDLQTKELQ KQITKIQKFR VYYRDSRDPI WKGPAKLLWK
D_CD_84ZR0   SAGERIIDII ASDIQTRELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
D_CD_ELI_K   SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
D_CD_NDK_M   SAGERIIDII ATDIQTRELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
D_UG_94UG1   SAGERIIDII ATDIQTKELQ KQIIKIQNFR VYYRDSRDPV WKGPAKLLWK
F1_BE_VI85   SAGERIIDII STDIQTRELQ KQITKIQNFR VYYRDSRNPV WKGPAKLLWK
F1_BR_93BR   SAGERTIDII ATDIQTRELQ KQIIKIQNFR VYYRDSRDPV WKGPAKLLWK
F1_FI_FIN9   SAGERIIDII ATDIQTKELQ KQVTKIQNFR VYYRDSRDPV WKGPAKLLWK
F1_FR_MP41   SAGERIIDII STDIQTRELQ KQIIKIQNFR VYYRDSRDPV WKGPAKLLWK
F2_CM_MP25   SAGERIIDII ATDIQTKELQ KQISKIQNFR VYFRDSRDPV WKGPAKLLWK
F2KU_BE_VI   SAGERIVDII ASDIQTRALQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
G_BE_DRCBL   SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
```

```
G_NG_92NG0   SAGERIIDII ASDIQTKELQ KQIIKIQNFR VYYRDSRDPI WKGPAKLLWK
G_SE_SE616   SAGERIIDII ASDIQTKELQ KQITKIQNFR VYYRDSRDPV WKGPAKLLWK
H_BE_VI991   SARERIIDII ATDIPTKELQ KQISQIQKFR VYYRDSRDPI WKGPAKLLWK
H_BE_VI997   SAGERIIDII ATDIQTKELQ KQISNIQKFR VYYRDSRDPI WKGPAKLLWK
H_CF_90CF0   SAGERIIDII ATDIQTKELQ KQISNIQKFR VYYRDSRDPI WKGPAKLLWK
J_SE_SE702   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
J_SE_SE788   SAGERIIDII ATDIQTRELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK
K_CD_EQTB1   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLPWK
K_CM_MP535   SAGERIVDII ATDIQTKELQ KQILNIQKFR VYYRDSREPI WKGPAKLLWK
N_CM_YBF30   TAGERIIDII ATDIQTTNLQ TQILKVQNFR VYYRDSRDPI WKGPAKLLWK
O_CM_ANT70   TAGERIIDIL ASQIQTTELQ KQILKXHKFR VYYRDSRDPI WKGPAQLLWK
O_CM_MVP51   TAGERLIDIL ASQIQTTELQ KQILKINNFR VYYRDSRDPI WKGPAQLLWK
O_SN_99SE_   TAGERIIDIL ASQIQTTELQ KQIFKIQKFQ VYYRDSRDPI WKGPAQLLWK
O_SN_99SE_   TAGERIIDIL ASQIQTTELQ KQIFKIQKFQ VYYRDSRDPI WKGPAQLLWK
U_CD___83C   SAGERIIDII ATDIQTKELQ KQITKIQNFR VYYRDSRDPI WKGPAKLLWK

             1001                                              1046
00BW0762_1   GE.GAVVIQD NSDIKVIPRR KAKIIKDYGK QMAGADCVAG RQDED.
00BW0768_2   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
00BW0874_2   GE.GAVVIQD NGDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
00BW1471_2   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGADCVAG RQDED.
00BW1616_2   GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGADCVAG RQDED.
00BW1686_8   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGADCVAG RQDEDQ
00BW1759_3   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDEDQ
00BW1773_2   GE.GAVVIQD NNDIKVVPRR KVKIIKGYGK QMAGADCVAG GQDEN.
00BW1783_5   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCMAG RQDEDQ
00BW1795_6   GE.GAVVLQD NSEIKVVPRR KVKIIRDYGK QMAGADCVAG RQDEDQ
00BW1811_3   GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGADCVAG GQDEN.
00BW1859_5   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGADCVAS RQDED.
00BW1880_2   GE.GAVVIQD KSDIKVVPRR KVKIIRDYGK QMAGADCVAD RQDED.
00BW1921_1   GE.GAVVIQD NSDVKVVPRR KAKIIRDYGK QMAGADCVAD RQDED.
00BW2036_1   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
00BW2063_6   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
00BW2087_2   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
00BW2127_2   GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAS RQDED.
00BW2128_3   GE.GAVVLQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG GQDEN.
00BW2276_7   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
00BW3819_3   GE.GAVVIQD NGDIKVVPRR KAKIIKDYGK QMAGADCVAS RQDEN.
00BW3842_8   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
00BW3871_3   GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
00BW3876_9   GE.GAVVIQD NSDIKVVPRR KAKIIKNYGK QMAGADCVAG RQDED.
00BW3886_8   GE.GAVVIQD KGDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
00BW3891_6   GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGDDCVAG RQDED.
00BW3970_2   GE.GAVVIQD NSDIKVVPRR RAKIIRDYGK QMAGADCVAD RQDED.
00BW5031_1   GE.GAVVIQD NSDIKAVPRR KAKIIKDYGQ QMAGADCVAG RQDEN.
 96BW01B21   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
  96BW0407   GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGDDCVAG RQDED.
  96BW0502   GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG GQDEN.
 96BW06_J4   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAS RQDED.
 96BW11_06   GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGADCVAG RQDED.
  96BW1210   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
 96BW15B03   GEGAVVVIQD NSDIKVVPRR KVKIIRDYGK QMAGADCVAG RQDED.
 96BW16_26   GE.GAVVLQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG GQDEN.
 96BW17A09   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGADCVAG RQDED.
 96BWMO1_5   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGADCVAG RQDEDQ
 96BWMO3_2   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGADCVAG RQDEDQ
98BWMC12_2   GD.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGADCVAG RQDED.
98BWMC13_4   GE.GAVVIQD NSEIKVVPRR KVKIIRDYGK QMAGADCVAG RQDEDQ
98BWMC14_a   GE.GAVVIQD SSDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
98BWMO14_1   GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
98BWMO18_d   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDEDQ
```

```
98BWMO36_a  GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG GQDED.
98BWMO37_d  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGADCVAG RQDEDQ
99BW3932_1  GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAS RQDED.
99BW4642_4  GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAD RQDED.
99BW4745_8  GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
99BW4754_7  GE.GAVVIQD KSDIKVVPRR KAKIIKDYGK QMAGDDCVAG RQDED.
99BWMC16_8  GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
A2_CD_97CD  GE.GAVVIQD NGDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
A2_CY_94CY  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
A2D___97KR  GE.GAVVIQD NSDIKVVPRR RAKIIRDYGK QMAGDDCVAG RQDED.
A2G_CD_97C  GE.GAVVIQD NNEIKVVPRR KTKILRDYGK QMAGDDCVAG RQDED.
A_BY_97BL0  GE.GAVVIQD NXDIKVVPRR KAKIIXDXXK QMAGXDCVAS RQDED.
A_KE_Q23_A  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
A_SE_SE659  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
A_SE_SE725  GE.GAVVIQD NNDIKVVPRR KAKILRDYGK QMAGDDCVAG RQDED.
A_SE_SE753  GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGDDCVAG RQDED.
A_SE_SE853  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
A_SE_SE889  GE.GAVLIQD NSDIKVVPRR KAKIIRDYGK QMAGDGCVAG RQDED.
A_SE_UGSE8  GE.GAVVIQD QSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
A_UG_92UG0  GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGDDCVAG RQDED.
A_UG_U455_  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCMAG RQDED.
AC_IN_2130  GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
AC_RW_92RW  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
AC_SE_SE94  GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGDDCVAG RQDED.
ACD_SE_SE8  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDEDW
ACG_BE_VI1  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDGCVAG RQDED.
AD_SE_SE69  GE.GAVVIQD NSEIKVVPRR KVKIIRDYGK QMAGDDCVAS RQDED.
AD_SE_SE71  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
ADHK_NO_97  GE.GAVVIQD NGDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
ADK_CD_MAL  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG GQDED.
AG_BE_VI11  GE.GAVAIQD NNEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
AG_NG_92NG  GE.GAVVIQD NSEIKVVPRR KVKIIKDYGK QMAGGDCVAG RQDED.
AGHU_GA_VI  GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
AGU_CD_Z32  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
AJ_BW_BW21  GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
B_AU_VH_AF  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_CN_RL42_  GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGDDCVAS RQDED.
B_DE_D31_U  GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
B_DE_HAN_U  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMGSDDCVAS RQDED.
B_FR_HXB2_  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_GA_OYI___ GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_GB_CAM1_  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_GB_GB8_A  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_GB_MANC_  GE.GAVVIQD NSEIKVVPRR KVKIIRDYGK QMAGDDCVAG RQDED.
B_KR_WK_AF  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_NL_3202A  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_TW_TWCYS  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
B_US_BC_L0  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_US_DH123  GE.GAVVIQD KSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_US_JRCSF  GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGDDCVAS RQDED.
B_US_MNCG_  GE.GAVVIQD NNDIKVVPRR KAKVIRDYGK QTAGDDCVAS RQDED.
B_US_P896_  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_US_RF_M1  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_US_SF2_K  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_US_WEAU1  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_US_WR27_  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
B_US_YU2_M  GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
BF1_BR_93B  GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGGDCVAG RQDED.
C_BR_92BR0  GE.GAVVLQD NSDIKVVPRR KVKIIKDYGK QMAGADCMAS RQDED.
C_BW_96BW0  GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGADCVAG RQDED.
C_BW_96BW1  GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGADCVAG RQDED.
```

```
C_BW_96BW1   GE.GAVVIQD NSDIKVVPRR KVKIIKDYGK QMAGADCVAG RQDED.
C_BW_96BW1   GEGAVVVIQD NSDIKVVPRR KVKIIRDYGK QMAGADCVAG RQDED.
C_ET_ETH22   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGADCVAG RQDED.
C_IN_93IN1   GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
C_IN_93IN9   GE.GAVVLQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
C_IN_93IN9   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGADCVAG RQDED.
C_IN_94IN1   GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
C_IN_95IN2   GE.GAVVIQD NSDIKVVPRR KAKIIKDYGK QMAGADCVAG RQDED.
CRF01_AE_C   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDEN.
CRF01_AE_C   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQNED.
CRF01_AE_C   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF01_AE_T   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF01_AE_T   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
CRF01_AE_T   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF01_AE_T   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF01_AE_T   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF01_AE_T   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF02_AG_F   GE.GAVVIQD KSDIKVVPRR KAKIIKDYGK QMAGDDCVAG RQDED.
CRF02_AG_F   GE.GAVVIQD KSDIKVVPRR KAKIIKDYGK QMAGDDCVAG RQDED.
CRF02_AG_G   GE.GAVVIQD NSDIKVVPRR KAKILRDYGK QMAGDDCVAG RQDED.
CRF02_AG_N   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF02_AG_S   GE.GAVVIQD NSDIKVVPRR KVKIVRDYGK QMAGDDCVAG RQDED.
CRF02_AG_S   GE.GAVVIQD NSDIKVVPRR KTKILRDYGK QMAGDDCVAG GQNED.
CRF03_AB_R   GE.GAVVIQD NNDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
CRF03_AB_R   GE.GAVVIQD NNDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
CRF04_cpx_   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGNDCVAG RQDED.
CRF04_cpx_   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF04_cpx_   GE.GAVVIQD NSDIKVVPRK KAKIIRDYGK QMAGDDCVAG RQDED.
CRF05_DF_B   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF05_DF_B   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF06_cpx_   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF06_cpx_   GE.GAVVIQD NSEIKVVPRR KAKIIKDYGK QMAGDDCVAG RQDED.
CRF06_cpx_   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF06_cpx_   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF11_cpx_   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
CRF11_cpx_   GE.GAIVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAG SQDED.
D_CD_84ZR0   GE.GAVVIQD NSDIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
D_CD_ELI_K   GE.GAVVIQD KSDIKVVPRR KVKIIRDYGK QMAGDDCVAS RQDED.
D_CD_NDK_M   GE.GAVVIQD NSDIKVVPRR KVKIIRDYGK QMAGDDCVAS RQDED.
D_UG_94UG1   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDED.
F1_BE_VI85   GE.GAVVIQD NSEIKIVPRR KAKIIRDYGK QMAVDDCVAG RQDED.
F1_BR_93BR   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
F1_FI_FIN9   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
F1_FR_MP41   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
F2_CM_MP25   GE.GAVVIQD NNEIKVIPRR KAKIIRDYGK QMAGDDCVAG RQDED.
F2KU_BE_VI   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
G_BE_DRCBL   GE.GAVVIQD NNEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
G_NG_92NG0   GE.GAVVIQD NNEIKVVPRR KAKILKDYGK QMAGGDCVAG RQDED.
G_SE_SE616   GE.GAVVIQD NNEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
H_BE_VI991   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
H_BE_VI997   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
H_CF_90CF0   GE.GAVVIQD NSEIKVVPRR EAKIIRDYGK QMAGDDCVAS RQDED.
J_SE_SE702   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
J_SE_SE788   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
K_CD_EQTB1   GE.GAVVIN. .SEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
K_CM_MP535   GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAG RQDED.
N_CM_YBF30   GE.GAVVIQD NGDIKVVPRR KAKIIRDYGK QMAGDGCVAS GQDENQ
O_CM_ANT70   GE.GAVVIQD KGDIKVVPRR KAKIIREYGK QMAGTDSMAS GQTESE
O_CM_MVP51   GE.GAVVIQD KGDIKVVPRR KAKIIRDYGK QMAGTDSMAN RQTESE
O_SN_99SE_   GE.GAVVIQD KGDIKVVPRR KAKIIRHYGK QMAGTDSMAS GQTESE
O_SN_99SE_   GE.GAVVIQD KGDIKVVPRR KAKIIRHYGK QMAGTDSMAS GQTESE
```

U_CD___83C GE.GAVVIQD NSEIKVVPRR KAKIIRDYGK QMAGDDCVAS RQDEN.

Table 15. HIV Rev Sequence Alignment

GCG Multiple Sequence File.
Written by Omiga 1.1

Name: 00BW0762_1	Len:	129	Check:	4903	Weight:	1.00
Name: 00BW0768_2	Len:	129	Check:	5102	Weight:	1.00
Name: 00BW0874_2	Len:	129	Check:	5815	Weight:	1.00
Name: 00BW1471_2	Len:	129	Check:	4144	Weight:	1.00
Name: 00BW1616_2	Len:	129	Check:	5298	Weight:	1.00
Name: 00BW1686_8	Len:	129	Check:	3871	Weight:	1.00
Name: 00BW1759_3	Len:	129	Check:	4976	Weight:	1.00
Name: 00BW1773_2	Len:	129	Check:	5775	Weight:	1.00
Name: 00BW1783_5	Len:	129	Check:	6142	Weight:	1.00
Name: 00BW1795_6	Len:	129	Check:	5055	Weight:	1.00
Name: 00BW1811_3	Len:	129	Check:	5804	Weight:	1.00
Name: 00BW1859_5	Len:	129	Check:	5252	Weight:	1.00
Name: 00BW1880_2	Len:	129	Check:	4995	Weight:	1.00
Name: 00BW1921_1	Len:	129	Check:	6482	Weight:	1.00
Name: 00BW2036_1	Len:	129	Check:	4770	Weight:	1.00
Name: 00BW2063_6	Len:	129	Check:	5384	Weight:	1.00
Name: 00BW2087_2	Len:	129	Check:	4848	Weight:	1.00
Name: 00BW2127_2	Len:	129	Check:	5783	Weight:	1.00
Name: 00BW2276_7	Len:	129	Check:	5364	Weight:	1.00
Name: 00BW3819_3	Len:	129	Check:	5712	Weight:	1.00
Name: 00BW3842_8	Len:	129	Check:	5586	Weight:	1.00
Name: 00BW3871_3	Len:	129	Check:	5299	Weight:	1.00
Name: 00BW3876_9	Len:	129	Check:	4423	Weight:	1.00
Name: 00BW3886_8	Len:	129	Check:	5415	Weight:	1.00
Name: 00BW3891_6	Len:	129	Check:	5426	Weight:	1.00
Name: 00BW3970_2	Len:	129	Check:	2613	Weight:	1.00
Name: 00BW5031_1	Len:	129	Check:	4597	Weight:	1.00
Name: 96BW01B21	Len:	129	Check:	5653	Weight:	1.00
Name: 96BW0407	Len:	129	Check:	4310	Weight:	1.00
Name: 96BW0502	Len:	129	Check:	4675	Weight:	1.00
Name: 96BW06_J4	Len:	129	Check:	5079	Weight:	1.00
Name: 96BW11_06	Len:	129	Check:	5939	Weight:	1.00
Name: 96BW1210	Len:	129	Check:	5666	Weight:	1.00
Name: 96BW15B03	Len:	129	Check:	5102	Weight:	1.00
Name: 96BW16_26	Len:	129	Check:	5675	Weight:	1.00
Name: 96BW17A09	Len:	129	Check:	2825	Weight:	1.00
Name: 96BWMO1_5	Len:	129	Check:	5636	Weight:	1.00
Name: 96BWMO3_2	Len:	129	Check:	6552	Weight:	1.00
Name: 98BWMC12_2	Len:	129	Check:	3043	Weight:	1.00
Name: 98BWMC13_4	Len:	129	Check:	5518	Weight:	1.00
Name: 98BWMC14_a	Len:	129	Check:	4358	Weight:	1.00
Name: 98BWMO14_1	Len:	129	Check:	7531	Weight:	1.00
Name: 98BWMO18_d	Len:	129	Check:	5291	Weight:	1.00
Name: 98BWMO36_a	Len:	129	Check:	6801	Weight:	1.00
Name: 98BWMO37_d	Len:	129	Check:	4790	Weight:	1.00
Name: 99BW3932_1	Len:	129	Check:	5736	Weight:	1.00
Name: 99BW4642_4	Len:	129	Check:	6464	Weight:	1.00
Name: 99BW4745_8	Len:	129	Check:	6181	Weight:	1.00
Name: 99BW4754_7	Len:	129	Check:	5182	Weight:	1.00
Name: 99BWMC16_8	Len:	129	Check:	4245	Weight:	1.00
Name: A2_CD_97CD	Len:	129	Check:	2625	Weight:	1.00
Name: A2_CY_94CY	Len:	129	Check:	4125	Weight:	1.00
Name: A2D___97KR	Len:	129	Check:	4114	Weight:	1.00
Name: A2G_CD_97C	Len:	129	Check:	1115	Weight:	1.00
Name: A_BY_97BL0	Len:	129	Check:	9470	Weight:	1.00
Name: A_KE_Q23_A	Len:	129	Check:	2684	Weight:	1.00
Name: A_SE_SE659	Len:	129	Check:	4659	Weight:	1.00

Name: A_SE_SE725	Len:	129	Check:	4491	Weight:	1.00
Name: A_SE_SE753	Len:	129	Check:	3636	Weight:	1.00
Name: A_SE_SE853	Len:	129	Check:	1862	Weight:	1.00
Name: A_SE_SE889	Len:	129	Check:	2798	Weight:	1.00
Name: A_SE_UGSE8	Len:	129	Check:	6865	Weight:	1.00
Name: A_UG_92UG0	Len:	129	Check:	4427	Weight:	1.00
Name: A_UG_U455_	Len:	129	Check:	3229	Weight:	1.00
Name: AC_IN_2130	Len:	129	Check:	5110	Weight:	1.00
Name: AC_RW_92RW	Len:	129	Check:	5015	Weight:	1.00
Name: AC_SE_SE94	Len:	129	Check:	7976	Weight:	1.00
Name: ACD_SE_SE8	Len:	129	Check:	2296	Weight:	1.00
Name: ACG_BE_VI1	Len:	129	Check:	3968	Weight:	1.00
Name: AD_SE_SE69	Len:	129	Check:	4558	Weight:	1.00
Name: AD_SE_SE71	Len:	129	Check:	2678	Weight:	1.00
Name: ADHK_NO_97	Len:	129	Check:	1890	Weight:	1.00
Name: ADK_CD_MAL	Len:	129	Check:	5260	Weight:	1.00
Name: AG_BE_VI11	Len:	129	Check:	4003	Weight:	1.00
Name: AG_NG_92NG	Len:	129	Check:	5027	Weight:	1.00
Name: AGHU_GA_VI	Len:	129	Check:	1978	Weight:	1.00
Name: AGU_CD_Z32	Len:	129	Check:	1958	Weight:	1.00
Name: AJ_BW_BW21	Len:	129	Check:	2263	Weight:	1.00
Name: B_AU_VH_AF	Len:	129	Check:	4074	Weight:	1.00
Name: B_CN_RL42_	Len:	129	Check:	4483	Weight:	1.00
Name: B_DE_D31_U	Len:	129	Check:	5079	Weight:	1.00
Name: B_DE_HAN_U	Len:	129	Check:	4550	Weight:	1.00
Name: B_FR_HXB2_	Len:	129	Check:	3649	Weight:	1.00
Name: B_GA_OYI_M	Len:	129	Check:	3334	Weight:	1.00
Name: B_GB_CAM1_	Len:	129	Check:	3865	Weight:	1.00
Name: B_GB_GB8_A	Len:	129	Check:	3083	Weight:	1.00
Name: B_GB_MANC_	Len:	129	Check:	5502	Weight:	1.00
Name: B_KR_WK_AF	Len:	129	Check:	4156	Weight:	1.00
Name: B_NL_3202A	Len:	129	Check:	3826	Weight:	1.00
Name: B_TW_TWCYS	Len:	129	Check:	3546	Weight:	1.00
Name: B_US_BC_L0	Len:	129	Check:	4674	Weight:	1.00
Name: B_US_DH123	Len:	129	Check:	4202	Weight:	1.00
Name: B_US_JRCSF	Len:	129	Check:	3217	Weight:	1.00
Name: B_US_MNCG_	Len:	129	Check:	3512	Weight:	1.00
Name: B_US_P896_	Len:	129	Check:	3297	Weight:	1.00
Name: B_US_RF_M1	Len:	129	Check:	5527	Weight:	1.00
Name: B_US_SF2_K	Len:	129	Check:	3616	Weight:	1.00
Name: B_US_WEAU1	Len:	129	Check:	4435	Weight:	1.00
Name: B_US_WR27_	Len:	129	Check:	812	Weight:	1.00
Name: B_US_YU2_M	Len:	129	Check:	4948	Weight:	1.00
Name: BF1_BR_93B	Len:	129	Check:	3645	Weight:	1.00
Name: C_BR_92BR0	Len:	129	Check:	4262	Weight:	1.00
Name: C_BW_96BW0	Len:	129	Check:	4323	Weight:	1.00
Name: C_BW_96BW1	Len:	129	Check:	3054	Weight:	1.00
Name: C_BW_96BW1	Len:	129	Check:	3900	Weight:	1.00
Name: C_BW_96BW1	Len:	129	Check:	4051	Weight:	1.00
Name: C_ET_ETH22	Len:	129	Check:	3843	Weight:	1.00
Name: C_IN_93IN1	Len:	129	Check:	2878	Weight:	1.00
Name: C_IN_93IN9	Len:	129	Check:	4499	Weight:	1.00
Name: C_IN_93IN9	Len:	129	Check:	3994	Weight:	1.00
Name: C_IN_94IN1	Len:	129	Check:	4362	Weight:	1.00
Name: C_IN_95IN2	Len:	129	Check:	3765	Weight:	1.00
Name: CRF01_AE_C	Len:	129	Check:	4444	Weight:	1.00
Name: CRF01_AE_C	Len:	129	Check:	3760	Weight:	1.00
Name: CRF01_AE_C	Len:	129	Check:	3562	Weight:	1.00
Name: CRF01_AE_T	Len:	129	Check:	5676	Weight:	1.00
Name: CRF01_AE_T	Len:	129	Check:	6090	Weight:	1.00
Name: CRF01_AE_T	Len:	129	Check:	6846	Weight:	1.00

```
Name: CRF01_AE_T       Len:   129  Check: 5393  Weight:   1.00
Name: CRF01_AE_T       Len:   129  Check: 6189  Weight:   1.00
Name: CRF01_AE_T       Len:   129  Check: 5202  Weight:   1.00
Name: CRF02_AG_F       Len:   129  Check: 5063  Weight:   1.00
Name: CRF02_AG_F       Len:   129  Check: 3731  Weight:   1.00
Name: CRF02_AG_G       Len:   129  Check: 2202  Weight:   1.00
Name: CRF02_AG_N       Len:   129  Check: 4873  Weight:   1.00
Name: CRF02_AG_S       Len:   129  Check: 3995  Weight:   1.00
Name: CRF02_AG_S       Len:   129  Check: 6502  Weight:   1.00
Name: CRF03_AB_R       Len:   129  Check: 2858  Weight:   1.00
Name: CRF03_AB_R       Len:   129  Check: 2808  Weight:   1.00
Name: CRF04_cpx_       Len:   129  Check: 3912  Weight:   1.00
Name: CRF04_cpx_       Len:   129  Check: 3700  Weight:   1.00
Name: CRF04_cpx_       Len:   129  Check: 3297  Weight:   1.00
Name: CRF05_DF_B       Len:   129  Check: 3974  Weight:   1.00
Name: CRF05_DF_B       Len:   129  Check: 4062  Weight:   1.00
Name: CRF06_cpx_       Len:   129  Check: 2954  Weight:   1.00
Name: CRF06_cpx_       Len:   129  Check: 1655  Weight:   1.00
Name: CRF06_cpx_       Len:   129  Check: 2327  Weight:   1.00
Name: CRF06_cpx_       Len:   129  Check: 2706  Weight:   1.00
Name: CRF11_cpx_       Len:   129  Check: 2064  Weight:   1.00
Name: CRF11_cpx_       Len:   129  Check: 1685  Weight:   1.00
Name: D_CD_84ZR0       Len:   129  Check: 4305  Weight:   1.00
Name: D_CD_ELI_K       Len:   129  Check: 4483  Weight:   1.00
Name: D_CD_NDK_M       Len:   129  Check: 3024  Weight:   1.00
Name: D_UG_94UG1       Len:   129  Check: 3298  Weight:   1.00
Name: F1_BE_VI85       Len:   129  Check: 2602  Weight:   1.00
Name: F1_BR_93BR       Len:   129  Check: 2572  Weight:   1.00
Name: F1_FI_FIN9       Len:   129  Check: 3253  Weight:   1.00
Name: F1_FR_MP41       Len:   129  Check: 2465  Weight:   1.00
Name: F2_CM_MP25       Len:   129  Check: 2231  Weight:   1.00
Name: F2KU_BE_VI       Len:   129  Check: 461   Weight:   1.00
Name: G_BE_DRCBL       Len:   129  Check: 3194  Weight:   1.00
Name: G_NG_92NG0       Len:   129  Check: 4325  Weight:   1.00
Name: G_SE_SE616       Len:   129  Check: 2614  Weight:   1.00
Name: H_BE_VI991       Len:   129  Check: 2347  Weight:   1.00
Name: H_BE_VI997       Len:   129  Check: 1680  Weight:   1.00
Name: H_CF_90CF0       Len:   129  Check: 2751  Weight:   1.00
Name: J_SE_SE702       Len:   129  Check: 2099  Weight:   1.00
Name: J_SE_SE788       Len:   129  Check: 2149  Weight:   1.00
Name: K_CD_EQTB1       Len:   129  Check: 3510  Weight:   1.00
Name: K_CM_MP535       Len:   129  Check: 2798  Weight:   1.00
Name: N_CM_YBF30       Len:   129  Check: 3973  Weight:   1.00
Name: O_CM_ANT70       Len:   129  Check: 9677  Weight:   1.00
Name: O_CM_MVP51       Len:   129  Check: 8852  Weight:   1.00
Name: O_SN_MP129       Len:   129  Check: 1678  Weight:   1.00
Name: O_SN_MP130       Len:   129  Check: 2242  Weight:   1.00
Name: U_CD___83C       Len:   129  Check: 9312  Weight:   1.00

                1                                                  50
00BW0762_1  MAGRSGD... NDDTLLQAVR IIKILYQSNP YPK.PEGTRQ ARRNRRRRWR
00BW0768_2  MAGRSEDS.. .DATLLQAVR IIKILYQSNP YPK.PEGTRQ ARKNRRRRRR
00BW0874_2  MAGRSGD... SDEALLQAVR IIKVLYQSNP YPK.PEGTRQ ARKNRRRRWR
00BW1471_2  MAGRSGD... SDEALLQAVR IIRILYQSNP YPKPEG.TRQ ARKNRRRRWR
00BW1616_2  MAGRSGDS.. .DEALLQAVR TIKILYQSNP YPE.PKGTRQ ARKNRRRRWR
00BW1686_8  MAGRSGDS.. .DEALLQAIK SIKILYQSNP YPE.PQGTRQ AQRNRRRRWR
00BW1759_3  MAGRSGD... NDEAVLQAIR IIKILYQSNP YPK.PRGTRQ AQKNRRRRWR
00BW1773_2  MAGRSGDS.. .DEALLQAVK IIKILYQSNP YPE.PKGTRQ ARKNRRRRWR
00BW1783_5  MAGRSGD... SDEAVLQAVR IIKILYQSNP YPK.PEGTRQ ARKNRRRRWR
00BW1795_6  MAGRSGD... GDAALLQAVR IIKILYQSNP YPK.PEGTRQ ARKNRRRRWR
00BW1811_3  MAGRSGD... SDEELLQVAR IIKILYQSNP YPE.PRGTRQ ARKNRRRRWR
```

```
00BW1859_5  MAGRSEDS..  .DAALLQAAK IIKIIYQSNP YPE.PKGTRQ ARRNRRRRWR
00BW1880_2  MAGRSGD...  NDEALLQAVR IIKILYQSNP FPK.PEGTRQ ARKNRRRRWR
00BW1921_1  MAGRSGD...  NDEALLQAVR IIKILYQSNP YPE.PQGTRQ ARKNRRRRWR
00BW2036_1  MAGRSEDS..  .DEALLQAIR LIKILYQSNP YPE.PKGTRQ ARKNRRRRWR
00BW2063_6  MAGRSGDN.D  ADAALLQAVK IIKILYQSNP YPK.PEGTRQ ARKNRRRRWR
00BW2087_2  MAGRSGD...  SDEALLQAVR IIKILYQSNP YPK.PEGTRQ ARKNRRRRWR
00BW2127_2  MAGRSGD...  NDEARLQVVK IIKILYQSNP YPK.PEGTRQ ARKNRRRRWR
00BW2276_7  MAGRSGD...  SDEALLQAVR IIKIIYQSNP YPK.PEGTRQ ARKNRRRRWR
00BW3819_3  MAGRSGD...  SDEDLLKAVR LIKILYQSNP YPK.PEGTRR AQRNRRRRWK
00BW3842_8  MAGRSEDS..  .DEALLRVVR IIKILYQSNP YPE.PKGTRQ AQRNRRRRWR
00BW3871_3  MAGRSGDS..  .DEALLQAIR TIKILYQSNP YPE.PKGTRQ ARKNRRRRWR
00BW3876_9  MAGRSGDS..  .DEALLHAVR TIKILYXSNP YPE.PKGTRQ ARKNRRRRWR
00BW3886_8  MAGRSGDS..  .DEALLQAVR IIKILYQSNP YPE.PKGTRQ ARKNRKRRWR
00BW3891_6  MAGRSGDS..  .DEALLQAVR IIKILYQSNP YPE.HQGTRQ ARKNRRRRWG
00BW3970_2  MAGRSGDS..  .DEALLQAVR IIKILYQSNP YPK.PEGTRQ ARKNRRRRWR
00BW5031_1  MAGRSGDN..  .DEALLQAVK IIKILYQSDP YPK.PEGTRQ ARKNRRRRWR
 96BW01B21  MAGRSGD...  SDEALLQAVR IIRILYQSNP YPK.PEGTRQ ARKNRRRRWR
  96BW0407  MAGRSGD...  SDEALLQAVR IIRILYQSNP YPE.PRGTRR ARKNRRRRWR
  96BW0502  MAGRSGDS..  .DEALLQAVK IIKILYQSNP YPK.PEEIRQ ARKNRRRRWR
 96BW06_J4  MAGRSGDS..  .DEALLQAVK AIKILYQSNP YPE.PKGTRQ ARKNRRRRWR
 96BW11_06  MAGRSGD...  NDEALLQAVR IIKILYQSNP SPE.PKGNRQ ARKNRRRRWR
  96BW1210  MAGRSGD...  SDEALLQAVR IIKILYQSNP YPK.PEGTRQ ARKNRRGRWR
 96BW15B03  MAGRSEDS..  .DEALLHAVR IIKILYQSNP YPE.PKGTRQ ARKNRRRRWR
 96BW16_26  MAGRSGDS..  .DAALLQAVR IIKILYQSNP YPK.PKGTRQ ARKNRRRRWR
 96BW17A09  MAGRSGD...  NDEALLQAMG IIKILYQSNP YPKPEG.TRR ARKNRRRRWR
 96BWMO1_5  MAGRSGD...  SDEALLQAVR IIKILYQSNP YPK.PEGTRQ ARRNRRRRWR
 96BWMO3_2  MAGRSGD...  SDEALLQAVR TIKILYQSNP YPK.PEGTRQ ARKNRRRRWR
98BWMC12_2  MAGRSGDS..  .DEALLQAVR IIKILYQSNP QPK.PEGTRQ ARKNRRRRWR
98BWMC13_4  MAGRSGD...  SDEALLQAVR IIKILYQSNS YPK.PEGTRQ ARKNRRRRWR
98BWMC14_a  MAGRSGDS..  .DEALLQAVR IIKILYQSNP PPE.RRGIGQ ARXNRRRRWR
98BWMO14_1  MAGRSGD...  DDERLLQAVR IIKILYQSNP YPS.PEGTRQ ARRNRRRRWR
98BWMO18_d  MAGRSGD...  SDEALLQAVR IIKILYQSNP YPK.PEGTRQ ARKNRRRRWR
98BWMO36_a  MAGRSGV...  SDEALLQAVK IIKILYQSNP YPNNPEGSRQ AQRNRRRRWR
98BWMO37_d  MAGRSGD...  SDEALLQAVR IIKILYQSNR YPK.PEGTRQ AQRNRRRRWR
99BW3932_1  MAGRSGD...  PDEALLQAIR IIKILYQSNP YPK.PEGTRQ ARRNRRRRWR
99BW4642_4  MAGRSEDSG.  .DAALLQAVR IIKILYQSNP YPE.PKGTRQ ARKNRRRRWR
99BW4745_8  MAGRSGDS..  .DEALLQAVR IIKILYQSNP YPK.PKETRQ ARRNRRRRWR
99BW4754_7  MAGRSGD...  NDAALLLAVQ TIKLLYQSNP YPK.PEGTRQ ARRNRRRRWR
99BWMC16_8  MAGRSGDS..  .DEALLQAVR IIKILYQSNP CPE.PRGTRQ ARKNRRRRWR
A2_CD_97CD  MAGRSGD...  PDEDLIRAIR IIKILYQSNP YPKPRG.TRQ ARKNRRRRWR
A2_CY_94CY  TAGRSDD...  PDESLLQAIR TIKILYQSNP YPKPRG.SRQ AQRNRRRRWR
A2D___97KR  MAGRSGD...  PDEDLLRAVR AIRILYQSNP SPDPRG.SRQ ARKNRRRRWR
A2G_CD_97C  MAGRSGS...  TDEELLQAAR IIKILYQSNP YPPPEG.TRQ ARKNRRRRWR
A_BY_97BL0  MAGRSGG...  TDAELLTAVR IIXFLYQSNP YPTPRK.TRQ AXKNQRRRXR
A_KE_Q23_A  MAGRSGD...  SDEELLRAVR IIKILYKSNP YPKPKG.SRQ ARKNRRRRWR
A_SE_SE659  MAGRSGP...  GDEELLKAIR IIKILYQSNP YPKPRG.TRQ ARKNRRRRWR
A_SE_SE725  MAGRSGD...  SDEELLRAVR IIKILYQSNP YPRPKG.SRQ AQKNRRRRWR
A_SE_SE753  MAGRSGN...  SDEELLRAIR IIKILYNSNP YPKPKG.SRQ ARKNRRRRWR
A_SE_SE853  MAGRSGN...  SDEELLRAIR IIKILYQSNP HPKPRG.SRQ ARKNRRRRWR
A_SE_SE889  MAGRSGD...  SDEELLKAVR IIKILYQSNP YPKPRG.TRQ ARKNRRRRWR
A_SE_UGSE8  MAGRSGD...  SDEELLKAVR TIKILYQSNP YPQPKG.SRQ ARKNRRRRWR
A_UG_92UG0  MAGRSGN...  PDEELLRAIR IIKILYQSNP YPEPKG.TRQ ARKNRRRRWR
A_UG_U455_  MARRSGN...  PDEDLLKAVR IIKLLYQSNP CPNPRG.SRQ ARKNRRRRWR
AC_IN_2130  MAGRSGD...  SDEELLQVVR IIKILYRSNP YPKPRG.TRQ ARKNRRRRWR
AC_RW_92RW  MAGRSGD...  SDETLLQAVK IIKILYQSNP YPK.PEGTRQ ARRNRRRRWR
AC_SE_SE94  MAGRSGD...  SDEALLQAVR IIKILYQSNP YPKPRG.TRQ ARKNRRRRWR
ACD_SE_SE8  MAGRSGD...  SDEDLLRAIR IIKILYKSNP YPKPRG.SRQ ARKNRRRRWR
ACG_BE_VI1  MAGRSGA...  SDEELLRAVR IVKILYQSNP YPKPEG.TRQ ARRNRRRRWR
AD_SE_SE69  MAGRSGD...  SDEXLLKAVR LIKTLYQSNP PPSPEG.TRQ ARRNRRRRWR
AD_SE_SE71  MAGRSGN...  SDEELLQAAR IIKILYQSNP YPKPKG.SRQ ARKNRRRRWR
ADHK_NO_97  MAGRSGD...  RDADLLKAVR IIKILYQSNP YPE.PTGSRQ ARRNRRRRWR
```

```
ADK_CD_MAL   MAGRSGD... SDEDLLRAIR LIKILYQSNP PPNTEGTTRQ ARRNRRRRWR
AG_BE_VI11   MAGRTGS... TDEELLKAVR TIKILYQSNP YPSSEG.SRQ ARKNRRRRWR
AG_NG_92NG   MAGRSGD... ADEELLRVTR IIKILYQSNP YPPPEG.TRQ ARKNRRRRWR
AGHU_GA_VI   MAGRSGA... SDEELLKAVR IIKILYQSNP FPE.PTGTRQ ARRNRRRRWR
AGU_CD_Z32   MAGRSGD... SDEELLKIVR IIKILYQGNP YPPPEG.TRQ ARRNRRRRWR
AJ_BW_BW21   MAGRSGD... NDEQLLLAIR IIKILYKSNP YPKPNG.SRQ ARRNRRRRWR
B_AU_VH_AF   MAGRSGD... SDDELLKTVR LIKFLYQSNP PPSPEG.TRQ ARRNRRRRWR
B_CN_RL42_   MAGRSED... SDEELLKTVR LIKLLYQSNP LPSPEG.TRQ ARRNRRRRWR
B_DE_D31_U   MAGRSGD... SDEELLKTVR LIKFLYQSNP PPSPEG.TRQ ARRNRRRRWR
B_DE_HAN_U   MAGRSGD... SDEELLKTVR LIKFLYQSNP PPSNEG.TPT ARRNRRRRWR
B_FR_HXB2_   MAGRSGD... SDEELIRTVR LIKLLYQSNP PPNPEG.TRQ ARRNRRRRWR
B_GA_OYI_M   MAGRSGD... SDEELLKTVR LIKFLYQSNP PPNPEG.TRQ ARRNRRRRWR
B_GB_CAM1_   MAGRSGD... SDEELLKAVR LIKLLYQSNP LPSSKG.TRQ ARRNRRRRWR
B_GB_GB8_A   MAGRSGD... SDEDFLKTVR FIKFLYQSNP PPNPKG.TRQ ARRNRRRRWR
B_GB_MANC_   MAGRSGD... SDEELLQTVK LVKFLYQSNP PPSPEG.TRQ ARRNRRRRWR
B_KR_WK_AF   MAGRSGD... SDEELLRTIR IIKFLYQSNP LPEPEG.TRQ ARRNRRRRWR
B_NL_3202A   MAGRSGD... SDEDLLKTVR LIKFLYQSNP PPSPEG.TRQ ARRNRRRRWR
B_TW_TWCYS   MAGRSGD... SDEELLRTVR LIKLIYQSNP PPNPEG.TRQ ARRNRRRRWR
B_US_BC_L0   MAGRSGD... SDEELLKTVR LIKLLYQSNP PPNPEG.TRQ ARRNRRRRWR
B_US_DH123   MAGRSGE... SDEDLLNTVR LIKLLYQSNP LPSLEG.TRQ ARRNRRRRWR
B_US_JRCSF   MAGRSGD... SDEDLLKTVR LIKFLYQSNP PPSNEG.TRQ ARRNRRRRWR
B_US_MNCG_   MAGRSGD... SDEELLKTVR LIKFLYQSNP PPSSEG.TRQ ARRNRRRRWR
B_US_P896_   MAGRSGD... SDEDLLKTVR LIKFLYQSNP PPSLEG.TRQ ARRNRRRRWR
B_US_RF_M1   MAGRRGD... SDEDLLKAVR LIKSLYQSNP PPSPEG.TRQ ARRNRRRRWR
B_US_SF2_K   MAGRSGD... SDEELLRTVR LIKLLYQSNP PPSPEG.TRQ ARRNRRRRWR
B_US_WEAU1   MAGRSGD... SDEDLLKTVR LIKILYQSNP PPSPEG.TRQ ARRNRRRRWR
B_US_WR27_   MAGRSGD... SDEELLQKV. LIRFLYQSNP PPSSEG.TRQ ARRNRRRRWE
B_US_YU2_M   MAGRSGD... SDEDLLRTVR LIKVLYQSNP PPSSEG.TRQ ARRNRRRRWR
BF1_BR_93B   MAGRSGD... SDTELLKAVS YIKILYQSNP LPKPKG.TRQ ARRNRRRRWR
C_BR_92BR0   MAGRSGD... SDEALLQAVR IIKILYQSNP YPK.PEGTRQ ARRNRRRRWR
C_BW_96BW0   MAGRSGD... SDEALLQAVR IIKILYQSNP YPK.PEGTRQ AWRNRRRRWR
C_BW_96BW1   MAGRSGD... NDEALLQAVR IIKILYQSNP YPK.PEGTRQ ARKNRRRRWR
C_BW_96BW1   MAGRSGD... SDEALLQAVR IIKILYQNNP YPK.PEGTRQ ARKNRRRRWR
C_BW_96BW1   MAGRSEDS.. .DEALLHAVR IIKILYQSNP YPE.PKGTRQ ARKNRRRRWR
C_ET_ETH22   MAGRSGD... SDEELLKAVR IIKILYQSNP YPT.PEGTRQ ARRNRRRRWR
C_IN_93IN1   MAGRSGDS.. .DEALLQAVR IIKILYQSNP YPE.PKGTRQ ARKNRRRRWR
C_IN_93IN9   MAGRSGDS.. .DEELLKAVR IIKILYQSNP YPE.PRGTRQ ARKNRRRRWR
C_IN_93IN9   MAGRSGDS.. .DEALLRAVR IIKILYQSNP YPE.PRGTRQ ARKNRRRRWR
C_IN_94IN1   MAGRSGDS.. .DEALLKAVR IIKILYQSNP YPE.PRGTRQ ARKNRRRRWR
C_IN_95IN2   MAGRSGDS.. .DEALLKAVR IIKILYQSNP YPE.PRGTRQ ARKNRRRRWR
CRF01_AE_C   MAGRSGN... TDEDLLQAVR IIKILYQSNP YPPPEG.TRQ ARKNRRRRWR
CRF01_AE_C   MAGRSGS... TDEDLLRTVR IIKILYQSNP YPPAEG.TRQ ARKNRRRKWR
CRF01_AE_C   MAGRSGS... TDEELLRAAR AIKILFQSNP YPSSEG.TRQ ARKNRRRRWR
CRF01_AE_T   MAGRSGS... TDEELLRAVR IINILYQSNP YPSSEG.TRQ TRKNRRRRWR
CRF01_AE_T   MAGRSGS... TDEELLRAVR IIKVLYQSNP YPSSEG.TRQ TRKNRRRRWR
CRF01_AE_T   MAGRSGS... TDEELLRAVR IINILYQSNP YPSSEGGTRQ TRKNRRRRWR
CRF01_AE_T   MAGRSGS... TDEELLRAVR IIKLLYESNP PPSSEG.TRQ TRKNRRRRWR
CRF01_AE_T   MAGRSGS... TDEELLRAVR IIKILYQSNP FPSSEG.TRQ TRRNRRRRWR
CRF01_AE_T   MAGRSGS... TDEELLRAVK IINILYQSNP LPSSEG.SRQ TRKNRRRRWR
CRF02_AG_F   MAGRSGD... ADEELLRVVR IIKILYQSNP YPPPEG.TRQ TRKNRRRRWR
CRF02_AG_F   MAGRSXD... ADEELLRVVR IVKILYQSNP YPPPEG.TRQ ARKNRRGRWR
CRF02_AG_G   MAGRSGD... ADEELLRVIR IIKILYQSNP YPKPEG.TRQ ARRNRRRRWR
CRF02_AG_N   MAGRSGD... ADEELLRAVR IIKILYQSNP YPPPEG.TRQ TRKNRRRRWR
CRF02_AG_S   MAGRSGN... ADEELLRAVR TIKILYQSNP YPPPEG.TRQ ARKNRRRRWR
CRF02_AG_S   MAGRSGD... ADEGLLRAVR IIRILYQSNP YPPPEG.SRQ ARRNRRRRWR
CRF03_AB_R   MAGRSGD... SDEELLKTIR LIKFLYQSNP PPSPEG.TRQ ARRNRRRRWR
CRF03_AB_R   MAGRSGD... SDEDLLKTIR LIKFLYQSNP PPNPEG.TRQ ARRNRRRRWR
CRF04_cpx_   MAGRSGN... IDEDLFKAAR AIKILYQSNP YPNNPTGTRQ ARRNRRRRWR
CRF04_cpx_   MAGRSG.... SNEDLLGSVG IVKILYQSNP YPN.PTGTRK ARRNRRRRWR
CRF04_cpx_   MAGRSGS... TDEDLLKAVG IVKILYQSNP YPNNTAGTRQ ARRNRRRRWR
CRF05_DF_B   MAGRSGD... RDEDLLKAVR LIKILYQSNP LPSPEG.TRQ ARRNRRRRWR
```

```
CRF05_DF_B  MAGRSGD... RDEDLLKAVR LIKFLYQSNP PPRPEG.TRQ ARRNRRRRWR
CRF06_cpx_  MAGRSGD... SDDRLLLAVR IIKILYQSNP YPKPNG.SRQ ARRNRRRRWR
CRF06_cpx_  MAGRSGD... NDEQLLLAVR IIKILYQNNP YPKPNG.GRQ ARRNRRRRWR
CRF06_cpx_  MAGRSGD... NDEQLLLAVR TIKILYQSNP YPKPSG.SRQ ARRNRRRRWR
CRF06_cpx_  MAGRSGD... SDEQLLWAVR VIKILYQSNP YPKLSG.SRQ ARRNRRRRWR
CRF11_cpx_  MAGRSGD... NDEQLLTAVK IIKILYQSNP QPNPTG.SRQ ARRNRRRRWR
CRF11_cpx_  MAGRSGD... SDAQLLAAAR IIKILYQSSP YPKPAG.TRQ ARRNQRRRWR
D_CD_84ZR0  MAGRSGD... SDEDLLTAVR LIKILYQSNP PPSPEG.TRQ ARRNRRRRWR
D_CD_ELI_K  MAGRSGD... SDEDLLKAVR LIKFLYQSNP PPSPEG.TRQ ARRNRRRRWR
D_CD_NDK_M  MAGRSGD... SDENLLKAIR LIKFLYQSNP PPSPEG.TRQ ARRNRRRRWR
D_UG_94UG1  MAGRSGD... RDEELLQAVR LIKILYQSNP PPSPEG.TRQ ARRNRRRRWR
F1_BE_VI85  MAGRSGD... SDTELLKAVK CIKILYQSNP YPKPEG.TRQ ARRNRRRRWR
F1_BR_93BR  MAGRSGD... SDQELLKAVR YIKILYQSNP YPKPEG.TRQ ARRNRRRRWR
F1_FI_FIN9  MAGRSGD... SDTELLKAVK YIKILYQSNP YPSPDG.TRQ ARRNRRRRWR
F1_FR_MP41  MAGRSGD... NDEELLRAVR AIKILYQSNP YPKPEG.TRQ ARRNRRRRWR
F2_CM_MP25  MAGRSGD... RDEELLKAVR YIKILYQSNP YPKLEG.TRK ARRNRRRRWR
F2KU_BE_VI  MAGRSGD... SDEELLKAVR LIKILYQSNP YPKPEG.TRQ ARRNRRKRWR
G_BE_DRCBL  MAGRSGS... TDEELLTAVR IIKLLYQSNP SPPPEG.TRQ ARRNRRRRWR
G_NG_92NG0  MAGRSGD... PDEELLRAVR IIKTLYQSNP YPSPAG.TRQ ARKNRRRRWR
G_SE_SE616  MAGRSGS... TDEELLRAVK AIKILYQSNP YPPPEG.TRQ ARRNRRRRWR
H_BE_VI991  MAGRSGD... NDEGLLRACR IIRLLYQSNP YPE.PAGTRQ AQRNRRRRWR
H_BE_VI997  MAGRSGA... GDEQLPQVCK IIKIIYQSNP YPE.PAGTRQ ARRNRRRRWR
H_CF_90CF0  MAGRSGA... SDTELLQVCK IIKILYQSNP CPE.PTGTRQ ARRNRRRRWR
J_SE_SE702  MAGRSGD... NDDQLLLAVR IIKILYQSNP YSKPNG.SRQ ARRNRRRRWR
J_SE_SE788  MAGRSGD... SDDQLLLAVR LIKILYQSNP YPKPNG.SRQ ARRNRRRRWR
K_CD_EQTB1  MAGRRGD... SEQQLLTPVR IIKILYQSNP YPKPEG.TRQ ARRNRRRRWR
K_CM_MP535  MAGRRGD... PDEQLLTTVR TIKILYESNP YPNLEG.SRQ TRRNRRRRWR
N_CM_YBF30  MAGRSGVN.. .DEELLRAVR VIKILYQSNP YPNSKG.TRQ ARRNRRRRWR
O_CM_ANT70  MAGRSED... .DQ.LLQAIQ IIKILYQSNP QPSPRG.SRN ARKNRRRRWR
O_CM_MVP51  MAGRSEE... .DQQLLQAIQ IIKILYQSNP CPTPAG.SRN ARKNRRRRWR
O_SN_MP129  MAGRSDG... .DQPLLRAIQ IIKILYQSNS HTSPTG.SRS ARRNRRRRWR
O_SN_MP130  MAGRSDG... .DQPLLRAIQ IIKILYQSNP HPSPTTGSTS ARRNRRRRWR
U_CD___83C  MAGRSGE... SDEELLRAVR IIKILYQSNP PPNPEG.TRQ ARKNRRRRWR

            51                                                  100
00BW0762_1  ARQRQIHSIS ERILSTVLGR PAEPVPFQLP PIERLHIGCS ESGGTSGTQQ
00BW0768_2  ARQRQINSIS ERILSTCLGR PAEAVPLQLP PIERLHIGCN ESGGTSGTQQ
00BW0874_2  ARQRQINSIS GRILSACLGR PTEPVPFQLP PIERLHINCS ENGGTSGTQQ
00BW1471_2  ARQRQIRAIS ERILDTCLGR LTEPVHLPLP PLERLHLDCS EDCEPTGTEQ
00BW1616_2  ARQRQIHSIS ERILSACLGR SAEPVPFQLP PIERLHIDCS ESSGNSG...
00BW1686_8  ARQRQIHSIS ERILSTCLGR SAEPVPLQLP PIERLHIDCS ESGGTFGTQQ
00BW1759_3  ARQRQIDSIS ERVLSTVLGR PTEPVPFQLP PIERLDIGDS ESGGTSGTER
00BW1773_2  ARQRQIREIS QRILSTYLGR PAEPVPLLLP PIERLHIDCS ESGGTSGTQQ
00BW1783_5  ARQRQIHSIS ERILSTCLGR STEPVPFQLP PIERLHIGDS KSSGTSGTQQ
00BW1795_6  ARQRQIHSIS ERILSTCLGR PAEPVPFLLP PLERLHIGDS ESSGTSGTQQ
00BW1811_3  ARQRQIREIS ERILCTCLGR STEPVPFLLP PIERLHIGDS EGSGTSGTQQ
00BW1859_5  ARQRQIHTIS ERILSTCLGR PAEPVPLQLP PIERLHIDCS ESSGTSGTQQ
00BW1880_2  ARQRQIHSIS ERILSTCLGR STEPVPFQLP PIERLHIGDS ESSGTSGTQR
00BW1921_1  ARQRQINSIS ERILTTCLGR SEEPVPLQLP PIERLNIGGS ESSGTSGTQQ
00BW2036_1  ARQRQIDSIS ARILSTCLGR PAEPVPFQLP PIERLNIGDN ESGGTSGTQQ
00BW2063_6  ARQKQIHSIS ERILSTCLGR SEEPVPLLLP PIERLRIGDS ESSGTSGTQQ
00BW2087_2  ARQREIHSIS ERILTTVLGR SAEPVPFQLP PIERLNINCS EGSGTSGTQQ
00BW2127_2  ARQRQIHSIS ERILSTCLGR STEPVPFQLP PIERLNIGDS EGSGTSGTQR
00BW2276_7  ARQRQIHTIS ERILSTCLGR PAEPVPLQLP PIERLYIGDS EGGGTSGTQQ
00BW3819_3  ARQRQIHSIS ERILSACLGR PAEPVPFQLP PIERLNIGDS EGDRTYGTQQ
00BW3842_8  ARQRQIDSIS GRLLSTCLGR SAEPVPLQLP PIERLNIGDS ESGGTSGTQQ
00BW3871_3  ARQRQINSIS ERILSTCLGR SAEPVPLQLP PIERLHIDCS ESGGTSGTQQ
00BW3876_9  ARQKQIHSIS ERILSACLGR SAEPVPLQLP PIERLHIGGS ESGGTSGTQQ
00BW3886_8  ARQRQINSIT ERILSDCLGR SAEPVPLQLP PIERLHIVDS ESGGTSGTQQ
00BW3891_6  ARQRQIHSLS ERILSTCLGR SAEPVPLQLP PLERLHIGDS ESGGTSGTQQ
00BW3970_2  ARQKEIHSLS ERILSTCLGR PAEPVPLQLP PLERLHIDCS ESGGTSGTQ.
```

```
00BW5031_1  ARQRQIHSIG ERILSTCLGR SAEPVPLQLP PLERLHIDCS ESSGTPGTQQ
 96BW01B21  ARQRQIHSIS ERILTTCLGR FTEPVPFQLP PIERLHIGDS ESGGTSGTQQ
  96BW0407  ARQRQIHSIS ERILSTCLGR PTEPVPFQLP PIERLHIDCS ESSGASGTQR
  96BW0502  ARQRQIHSLS ERILSACLGR PAEPVPFQLP PIERLHIDCS KSVGTSG...
 96BW06_J4  ARQRQINSIS ERILSTCLGR SAEPVPFQLP PIERLHIGGS ESGGTSGTQQ
 96BW11_06  TRQKQINSIS ERILSTCLGR SAEPVPFLLP PIERLHISDS ESSGTSGTQQ
  96BW1210  ARQRQIHSIS ERILSTCLGR PAEPVPLQLP PIERLHIGGS ENSGTTGTQQ
 96BW15B03  ARQRQIDSIS TRILSTCLGR PEEPVPFQLP PIERLNIGDS ESGGTSGTQQ
 96BW16_26  ARQRQIRAIS ARIFSACLGR PAEPVPLQLP PIERLHIGSS ESGGTSGTQ.
 96BW17A09  ARQRQIHSIS ERILSTCLGR PTEPVHLPLP PLERLHLDCS EDCEPTGTEQ
 96BWMO1_5  ARQRQINSIS ERILSACLGR PAEPVPFLLP PIERLHIGDS ESSGTSGTQQ
 96BWMO3_2  ARQRQIHSIS ERILSACLGR PTEPVPFQLP PIERLSIGDS ESSGTSGTQQ
98BWMC12_2  ARQKHIHSIS ERILSDCLGR SAEPVPLQLP PIERLHIDCS ESGGTSGTQQ
98BWMC13_4  ARQRQINSIS ERILSTCLGR STEPVPFLLP PIERLHIGDS ESSGTSGTQQ
98BWMC14_a  ARQTQINSLS ERIVSTCLGR PAERAPFQLP PIEKLHIDCS ESGGTPG...
98BWMO14_1  ARQRQIRALS ERILSTVLGR PTEPVPFQLP PIERLNINCS ESSGTSGTQQ
98BWMO18_d  ARQRQINSIS ERILSTCLGR PAEPVPFQLP PIERLHIDCS ESSGTSGTQH
98BWMO36_a  ARQRQIHSIS ERILSSCLGR PTEPVPFQLP PIERLHIGDS ESGGTYGTQQ
98BWMO37_d  ARQRQIHSIS ERILSTCLGR SAEPVPLQLP PIERLHIADS KGGGTSGIQQ
99BW3932_1  ARQRQIHSIS ERILFTCLGR STEPVPFQLP PIERLNIGDS ESGGTYGTLQ
99BW4642_4  ARQRQISSLS ERILSTCLGR SAEPVPLQLP PIERLHIDCS ESSGTSGTQQ
99BW4745_8  ARQNQIRAIS KRILSTCLGR PAEPVPLQLP PLERLHIGDS ESGGTSGTQQ
99BW4754_7  ARQRQIHSIS ERILSNCLGR PAEPVPFQLP PIEGLHIDCN ESSGTSGTQQ
99BWMC16_8  ARQRQIDSIS QRILSDCLGG PAEPVSFQLP PIERLNIDCN ESGGTSGTQQ
A2_CD_97CD  ARQRQIDSIS ERILSTCLGR PTEPVPLQLP PLERLHLDCS EDCGTCGTQQ
A2_CY_94CY  ARQRQIDSIS ERVLRTCLGR PTEPVPLQLP PLERLHIDCS EDCGTSGTLQ
A2D___97KR  ARQRQIRAIS ERILSACLGR PTEPVPLQLP PLERLHLDCS EDCGTSGTQQ
A2G_CD_97C  ARQRQIREIS ERILSTCLGR PTEPVPLLLP PLERLHLDCS EDGGTAETQQ
A_BY_97BL0  ARQRQINSIS KRIFSTCXGR SXEPVSLQLP PIERLHLDCS EDCXTSETQQ
A_KE_Q23_A  ARQRQIDSIS ERILSTCLGR PTEPVPLQLP PLERLHLDCC EDCGTSGTQQ
A_SE_SE659  ARQRQIDSIS ERILSACLGR STEPVPLQLP PLERLNLDCR EDCGTSGTQQ
A_SE_SE725  ARQRQIDSIS ERILSTCLGR SEEPVPLQLP SLETLHLDCH DDCGTSGTQQ
A_SE_SE753  ARQRQIDSIS ERILNACLGG STEPVPLQLP PLERLNLDCR EDCGTSGTQQ
A_SE_SE853  ARQRQIDSIS ERILSTCLGR SAEPVPLQLP PLEGLHLDCC EDCGTSGTEG
A_SE_SE889  ARQKQIDSLS ERILSTCLGR PQEPVPLLLP PIERLHLDCS EDCGTSGTQQ
A_SE_UGSE8  ARQRQIDSLS QRILSACLGR SEEPVPLQLP PLERLHIDCH EDCGTSGTG.
A_UG_92UG0  ARQRQIDTLS ERVLSTCLGR PAEPVPLQLP PIERLHLDCS EDCGTSGTQQ
A_UG_U455_  ARQRQIDSLS ERILSDCLGR PAEPVPLQLP PIERLRLDCS ESCGTSGTQQ
AC_IN_2130  ARQRQIDSIS ERILSTFLGR SAEPVPLQLP PLERLHLDCQ EDCGTSGTQQ
AC_RW_92RW  ARQRQIHSIS ERILSTCLGR PTEPVPFQLP PIERLTIDCS EDGGTSGTQQ
AC_SE_SE94  ARQRQIDSIS ERILSTCLGR SAEPVPLQLP PLERLHLD.. .....SGTQQ
ACD_SE_SE8  ARQRQIDSIS QRILSTCLGR SEEPVPLQLP PLERLNLDCC EDCGTSGTQG
ACG_BE_VI1  ARQRHIHSLS ERILCTCLGR SEEPVHLPLP PLEGLTLDCN ESSGTSGTEG
AD_SE_SE69  ARQRQINSIG ERILSTYLGR SQEPVPLQLP PLERLTLNCI EDCGTSGTQG
AD_SE_SE71  ARQNQIDSIS KRILSNCLGR PAEPVPLQLP PLERLNLNCS KDCGTSGTQG
ADHK_NO_97  ARQXQIHSIG ERVLATCMGR PAEPVPLQLP PLERLTLDSS EDCDIAGKQG
ADK_CD_MAL  ARQRQINSIG ERILSTYLGR PEEPVPLQLP PLERLTLNCN EDCGTSGTQG
AG_BE_VI11  ARQRHIQAIS RRILDACLGR PAEPVPLQLP PLERLSLDCS KDIGTSGTQR
AG_NG_92NG  ARQRQISALS ERILSTCLGR PAEPVPLQLP PIERLSLDCS EDSRTPETQQ
AGHU_GA_VI  ARQKQIHSIG ERVLATYLGR PAEPVPLQLP PLERLTLDCS EDCGTSGEKG
AGU_CD_Z32  ARQRQIHSLG ERILTTCLGR STEPVPFLLP PIERLRIDCS EDRGDSDPQG
AJ_BW_BW21  ARQNQIDSIS ERILSTCLGR PTEPVPFQLP PIERLRLDCS EDCGHSGTQG
B_AU_VH_AF  ARQRQIRQIS GWILSTYLGR PAEPVPLQLP PLERLTLDCS KDCGTSGTQG
B_CN_RL42_  ARQRQIREIS DRILVTYLGG STEPVPLQLP PLERLTLDCS KDCGTSGTQG
B_DE_D31_U  QRQRQIQSIS ERILSTYLGR PEQPVPLPLP PLERLTLDCS EDCGTSGTQG
B_DE_HAN_U  ERQRQIRSIS ERILSTFLGR PAEPVPLQLP PLERLTLDCS EDCGNSGTQG
B_FR_HXB2_  ERQRQIHSIS ERILGTYLGR SAEPVPLQLP PLERLTLDCN EDCGTSGTQG
B_GA_OYI_M  ERQRQIRKIS GWILSTYLGR SAEPVPLQLP PLERLNLDCS EDCGTSGTQG
B_GB_CAM1_  ERQRHIRAIS NWILSTHLGR PAEPVPLQLP PLERLTLDCS KDCGTSGTQG
B_GB_GB8_A  ARQRQIHQIG EWILSAFLGR PAEPVPLQLP PIERLTLDCD EDCGTSGTQG
B_GB_MANC_  GRQRQIQSLS AWILSTRLGR STQPVPLQLP PLERLTLDCS EDCGTSGTQG
```

```
B_KR_WK_AF  RRQWWIQSLS GWILNTHLGR PAEPVPLQLP PLERLTLDCN EECGTSGTQG
B_NL_3202A  ERQRQIRSIS ERILSTYLGR SAEPVPLQLP PLERLTLDCD EDCGTSGTQG
B_TW_TWCYS  ERQRQIRTIS GWILSNYLGR PAEPVPLQLP PLERLTLDCD EDCGTSGTQG
B_US_BC_L0  ERQRQIRSIS ERILSTFLGR SAEPVPLQLP PLERLNLGCN EDCGTSGTQG
B_US_DH123  QRQRQIQSIS GWILSNHLGR PADAVPLQLP PLERLTLDCN EDCGTSGTQG
B_US_JRCSF  ERQRQIRTIS ERILSTYLGR PAEPVPLQLP PLERLTLDCN EDCGTSGTQG
B_US_MNCG_  ERQRHIRSIS AWILSNYLGR PAEPVPLQLP PLERLTLDCN EDCGTSGTQG
B_US_P896_  ERQRQIRSIS ERILGTFLGR FEEPVPLPLP PLEKLTLDCN EDCGTSGTQG
B_US_RF_M1  ERQRQIRRCS EWILDTYLGR SVDPVQLQLP PLERLTLDSS EDCGTSGTQG
B_US_SF2_K  ERQRQIRSIS GWILSTYLGR SAEPVPLQLP PLERLTLDCS EDCGNSGAQG
B_US_WEAU1  ERQRQIRKIS GWILNTYLGR PTEPVPLPLP PLDRLTLDCK EDCGTSGTQG
B_US_WR27_  .RQRQIQSLS AWIISTHLGR PAEPVPLQLP PLERLTLDCS EDCGTSGTQG
B_US_YU2_M  ERQRQIRSIS GWLLSNYLGR PTEPVPFQLP PLERLTLDCN EDCGTSGTQG
BF1_BR_93B  ARQRQIREIS ERILSSCLGR PEEPVPLQLP PLERLHINCS EDCGQGTEEG
C_BR_92BR0  ARQRQIHSIS ERILSTCVGR PAEPVPFQLP PIERLNINCS ESGGTSGTQQ
C_BW_96BW0  ARQRQIHSIS ERILSTCLGR PTEPVPLQLP PIERLHIDCS ESSGASGTQQ
C_BW_96BW1  ARQKQINSIS ERILSTCLGR SAEPVPFLLP PIERLHISDS ESGGTSGTQQ
C_BW_96BW1  ARQRQIHSIS ERILSTCLGR PAEPVPLQLP PIERLHIGGS ENSGTTGTQQ
C_BW_96BW1  ARQRQIDSIS TRILSTCLGR PEEPVPFQLP PIERLNIGDS ESGGTSGTQQ
C_ET_ETH22  ARQRQIHTLS ERILSNFLGR PAEPVPLQLP PLERLNLDCS EDSGTSGTQQ
C_IN_93IN1  ARQRQIHSIS ERILSTCLGR STEPVPLQLP PIERLHIGGS ESGGTSGTQQ
C_IN_93IN9  ARQRQIHSLS ERILSACLGR PAEPVPLQLP PLERLHISGS ESGGTSGTQQ
C_IN_93IN9  ARQKQIHSLS ERILSTCLGR SAEPVPLQLP PLERLHISGS ESGGTSGTQQ
C_IN_94IN1  ARQRQIHSIS ERILSACLGR PAEPVPLQLP PIERLHISGS ESGGTSGTQQ
C_IN_95IN2  ARQRQIHSIS ERILSTFLGR PAEPVPLQLP PIERLHISGS ESAGTSGTPQ
CRF01_AE_C  RRQRQIHSLS ERILVACVGR STEPVPLQLP PLERLHIDCS EDCGTSGTQQ
CRF01_AE_C  ARQRQIHKIG ERILSTCLGR SPEPVPLQLP PLERLHLDCS EDCGTSGTQQ
CRF01_AE_C  ARQRQIRALS ERILSACLGR SAEPVPLQLP PLERLHLDCS EDCGTSGTQQ
CRF01_AE_T  ARQRQIRAIS ERILITCLGR STEPVPLQLP PLERLHLDCN EDCGTSGTQQ
CRF01_AE_T  ARQRQIRAIS ERILNACVGR STEPVPLQLP PLERLHLDCS EDCGTSGTQQ
CRF01_AE_T  ARQRQIRAIS ERILSTCLGR STEPVPLQLP PLERLHLDCS EDCGTSGTQQ
CRF01_AE_T  ARQRQIREIS ERILSSCVGR STEPVPLPLP PLERLHLDCS EDCGTSGTQQ
CRF01_AE_T  ARQRQISAIS ERILSTCLGR STEPVPLQLP PVERLNLDCS EDGGTSGTQQ
CRF01_AE_T  ARQRQISAIS ERILSACLGR STEPVSLPLP PLERLHLDCS EDCGTSGTQQ
CRF02_AG_F  ARQRQIRAIS ERFLSTCLGR SAEPVPLQLP PIERLCLDCS EGCGTSGTQQ
CRF02_AG_F  ARQRQIRAIS QRILSTCLGR SAEPVPLQLP PLERLCLDCS EGCGTSGTQQ
CRF02_AG_G  ARQRQIHSLS ERILSTCLGR PEEPVSFQLP PLERLNLDCS EDCGNSGTQS
CRF02_AG_N  ARQRQIRAIS ERILSTCLGR SAEPVPLQLP PIERLNLDCS EDCGTSGTQL
CRF02_AG_S  ARQRQIRAIS ERILSTCLGR SAEPVPLQLP PIERLRLDCS EDCGTSGTQG
CRF02_AG_S  ARQRQVRAIS ERILSTCLGR PAEPVPLPLP PIERLCLDCS EDSGTSGTQQ
CRF03_AB_R  ERQRHIHSIS EQILSTYLGR PEEPVLLHLP PLERLTLDCS EDCGTSGTQG
CRF03_AB_R  ERQRHIHSIS QRILSTYLGR PEEPVPLHLP PLERLTLDCS EDCGTSGTQG
CRF04_cpx_  ARQKQIHSLS ERILATYLGR PAEPVPLQLP PLEKLTLNCS EDCGTSGDKG
CRF04_cpx_  ARQKQIHSIS ERVLATYLGR PAEPVPLQLP PLEKLTLNCS EDCGTSGEKG
CRF04_cpx_  ARQNRIHSIS ERILAACLGR PAEPVPLQLP PIEKLTLDCS EDCGTSGDKG
CRF05_DF_B  ARQRQINSIG ERLLSTYLGR SEEPVPLQLP PLERLNLNCS EDCGTSGTQG
CRF05_DF_B  ARQRQIRSIA DRIVDTYLGR PEEPVPLQLP PLERLNLNCS EDCGTSGTQG
CRF06_cpx_  ARQNQIDSIS ERVLSTCLGR SAEPVPLQLP PIERLRLDCS EDCGNSGTQG
CRF06_cpx_  ARQNQIDSIS ERILSTCLGR PTEPVPFQLP PIERLRLDCS EDCGNSGTQG
CRF06_cpx_  ARQKQIDSIS ERILSTCLGR SAEPVPLQLP PIERLRLDCS EDCGNSGTQG
CRF06_cpx_  ARQNQIDSIS ERILSSCLGR SEEPVPLQLP PIERLRLDCT EDCGNSGTQG
CRF11_cpx_  ARQNQIDSIS QRILSDCLGR SEEPVPLQLP PIERLHLDCS EDCGNPGTQG
CRF11_cpx_  ARQNQLHSIS QRILSTCLGR SEEPVPLPLP PIERLHLDCS EDCGNSGTQG
D_CD_84ZR0  ARQRYIHSIG ERILSTYLGR SEEPVPLQLP PLERINLNCS EDCGTSGTQG
D_CD_ELI_K  ARQRQIREIA ERILGTYLGR PAEPVPLQLP PLERLNLNCS EDCRTSGTQG
D_CD_NDK_M  ARQRQIHSIG ERIICTFLGR PEEPVPLQLP PLERLNLNCS EDCGTSGTQG
D_UG_94UG1  ARQRQIHSIG ERIISTYLGR FEEPVPLQLP PLERLNLNCS EDCGTSGTQG
F1_BE_VI85  ARQRQIRALS DRILSSCLGR SEEPVPLQLP PLERLHINCS EDCGQGPEEG
F1_BR_93BR  ARQRQIREIS DRILSSCLGR PAEPVPLQLP PLERLHINCS EDCGQGAEEG
F1_FI_FIN9  ARQRQIRAIS ERILSSCLGR LEEPVPLQLP PLERLHINCS EDCGQGTEEG
F1_FR_MP41  ARQKQIRSIS ERILVACLGR PEEPVPLQLP PLERLHINCS KDCGQGTNEG
```

```
F2_CM_MP25   ARQRQIHQIS ERILSTCLGR LQEPVRLQLP LLEKLHINCS EDCGQGTEKG
F2KU_BE_VI   ARQRQIHSIS QRILSTCLGR PAEPVPFQLP PLERLNLDCS EDSREGAEGE
G_BE_DRCBL   ARQRQIHSIS ERILSTCLGR PEEPVPLQLP PLERLHLDCS EDGGTSGTQQ
G_NG_92NG0   ARQRQIHSIS ERILSACLGR PAEPVPFQLP PLEGLSLDCS KDGGTSGTQQ
G_SE_SE616   ARQRQISAIS ERILTAYLGR PAEPVPLQLP PLERLHLDCS EDSGTSGTQQ
H_BE_VI991   ARQRQIHSIG ERVLATCLGG PAEPVPLQLP PLERLTLDCS EDCGTSGEKG
H_BE_VI997   ARQRQIRAIS ERILTDCLGR PPEPVPLQLP PLERLTLDCN KDCGTSGEKG
H_CF_90CF0   ARQRQIREIS ERILTSCLGR PPEPVTLQLP PLERLTLNCS EDCGTSGEKG
J_SE_SE702   ARQNQIDSIS ERILSSCLGR PAEPVPLQLP PIERLRLDCS EDCGNSGTQG
J_SE_SE788   ARQNQIDSIS ERIPSSCLGR PAEPVPLQLP PIERLRLDCS EDCGNSGTQG
K_CD_EQTB1   ARQRQIREIS QRVLSSCLGR STEPVPLQLP PLERLSLNCD EDSGQGTEGE
K_CM_MP535   ARQKQISSIS ERLLSACLGR SAEPVPLQLP PIEKLNLNCD EDPGKGTEGG
N_CM_YBF30   ARQRQIRAIS ERILSSCLGG PPEPVDLPLP PLDRLTLDTE EDSGTPGTES
O_CM_ANT70   RRQAQVDTLA ARVLATVVHG PQNNNIVDLP PLEQLSIRDP EGDQLSEAWT
O_CM_MVP51   RRQAQVDSLA TRILATVVHG SQDNNLVDLP PLEQLNIRDP EADRLPGTGT
O_SN_MP129   TRHAHVDTLA ARILATVVHG PQDNNLVELP PLEQLSIRDP DGDQPSGTWT
O_SN_MP130   KRQAQIDTLA ARILATVVHG PQDNNLVELP PLEQLSIRDP DGDQPSGTWT
U_CD___83C   RRQQQIRSIS ERILSTCLGR PAEPVHLQLP PLERLNLDCS ....KGTATG

             101                              129
00BW0762_1   PQGTPEGMGN P......... .........
00BW0768_2   SQGTSEGVGS P......... .........
00BW0874_2   SQGTTEGVGN P......... .........
00BW1471_2   SQGITEGVGS P......... .........
00BW1616_2   ....TQGVGS P......... .........
00BW1686_8   SQGATEGVGN P......... .........
00BW1759_3   .......VGS P......... .........
00BW1773_2   SQGTTEGVGS P......... .........
00BW1783_5   SQGTTEGVGN P......... .........
00BW1795_6   SQGTPEGVGN P......... .........
00BW1811_3   SQGTPEGVGN P......... .........
00BW1859_5   SQGTTEGVGS P......... .........
00BW1880_2   SQGTPEGVGN P......... .........
00BW1921_1   SQGTTEGVGN P......... .........
00BW2036_1   SQGTTEGVGS P......... .........
00BW2063_6   SQGTPEGVGN P......... .........
00BW2087_2   PQGTTEGVGN P......... .........
00BW2127_2   .......VGS P......... .........
00BW2276_7   SQGTTEGVGS P......... .........
00BW3819_3   SQGTTEGVGS P......... .........
00BW3842_8   PQGTTEGVGS P......... .........
00BW3871_3   SQGTTEGVGN P......... .........
00BW3876_9   SQGTKEGVGS P......... .........
00BW3886_8   SQGTTEGVGS P......... .........
00BW3891_6   SQGTTEGVGS P......... .........
00BW3970_2   ......GVGH P......... .........
00BW5031_1   PQGDTEGVGR P......... .........
 96BW01B21   SQGTTEGVGN P......... .........
  96BW0407   SQGTTEGVGN P......... .........
  96BW0502   ....TEGVGS P......... .........
 96BW06_J4   SQGPTEGVGS P......... .........
 96BW11_06   SQGTPEGVGN P......... .........
  96BW1210   SQGTTEGVGS P......... .........
 96BW15B03   SQGTTEGVGS P......... .........
 96BW16_26   ......GVGS P......... .........
 96BW17A09   SQGATEGVGS P......... .........
 96BWMO1_5   SQGTPEGVGN P......... .........
 96BWMO3_2   SQGTTEGVGS S......... .........
98BWMC12_2   SQGTAEGVGS P......... .........
98BWMC13_4   SHGTPEGVGN P......... .........
98BWMC14_a   ....TQGVGN P......... .........
```

```
98BWMO14_1  SLGTTEGVGS P......... .........
98BWMO18_d  SQGTTEGVGN P......... .........
98BWMO36_a  PQGTTEGVGN P......... .........
98BWMO37_d  PQGTTEGVGS P......... .........
99BW3932_1  SQGTTEGVGS P......... .........
99BW4642_4  SQGTTEGVGS P......... .........
99BW4745_8  SQGTTEGVGS P......... .........
99BW4754_7  SQGTPEGVGN S......... .........
99BWMC16_8  SQGTTEGVGS P......... .........
A2_CD_97CD  SQGAETGVGR PQTSVESSGI LGSGIEDX.
A2_CY_94CY  SQGTETGVGR SQESVESSVI LGSGTEEX.
A2D___97KR  PQGTETGVGR PQISVEPSVV LGSGTEEX.
A2G_CD_97C  PQGTETGVGG .TIFVESSVI LGSRTKEQX
A_BY_97BL0  SQXTETXVXX PQISXESSXI XXSGTKEX.
A_KE_Q23_A  SQGAETGVGR HQVSVESPVI LGSGTKNX.
A_SE_SE659  SQGVETGVGR PQVSGESPVI LGSGTKNX.
A_SE_SE725  SQGVETGVGR PQVPGEPSTV LGSGTKTX.
A_SE_SE753  SQGIETGVGR PQVSVESPVI LGSGTKEX.
A_SE_SE853  .......VGR PQVSVESPGV LDSGTKNX.
A_SE_SE889  SQGAETGVGG PQVSEESSII LGSGTKTX.
A_SE_UGSE8  .......... TQVSGESSVV LDSGTKDX.
A_UG_92UG0  SQGVETGVGR TQVSGESPVV LGSGTKNX.
A_UG_U455_  PQGTETGVGG PQISVESSAV LGSGTKNX.
AC_IN_2130  SQGVETGVGR PQVSVESPGI LGSGTKNX.
AC_RW_92RW  SQGTTEGVGN PVSRKSCAVL GSGTKKEX.
AC_SE_SE94  SQGTETGVGR PQVSVESSAI LGPGTKNX.
ACD_SE_SE8  .......VGS NQISVESPAV LDSGTKEX.
ACG_BE_VI1  .......VGS SQTSGEHPVI LESGTKEX.
AD_SE_SE69  .......VGS PQIPVEPPAV LDSGTKEX.
AD_SE_SE71  .......VGS PQIPVESPAI LDSGTENX.
ADHK_NO_97  .......VGD PQIPGESSAV LGTGTKEX.
ADK_CD_MAL  .......VGS PQISVESPAI LGSGTEEX.
AG_BE_VI11  SQGTETGVGR PQIFVESSGV LGSGTKEX.
AG_NG_92NG  SPGTETGVGG PQISVESPVV LGSGTKEX.
AGHU_GA_VI  .......VGS PQISVESPTV LGTGAKEX.
AGU_CD_Z32  .......VGD SQIPGESCDL LGSGTKEX.
AJ_BW_BW21  .......VGD PQVSGESCPI LGEGTKEX.
B_AU_VH_AF  .......VGG PQVLVESPAV LESGAAEX.
B_CN_RL42_  .......VGS PQILVESPAV LDSGTKEX.
B_DE_D31_U  .......VGS PQILVESPAV LESGTKEX.
B_DE_HAN_U  .......VGS PQVLVESPAV LEPGTKEX.
B_FR_HXB2_  .......VGS PQILVESPTV LESGTKEX.
B_GA_OYI_M  .......VGS PEILVESPAV LEPGTKEX.
B_GB_CAM1_  .......VGS PQILVESPAV LESGTKEX.
B_GB_GB8_A  .......VGS PQVLVESPAV LDPGTKEX.
B_GB_MANC_  .......VGN PQVLVESPAV LESGSKEX.
B_KR_WK_AF  .......VGN PQILVESPAV LESGTKEX.
B_NL_3202A  .......VGS PQILVESPAV LESGTKEX.
B_TW_TWCYS  .......VGS PQIFVESPTV LDSGTKEX.
B_US_BC_L0  .......VGS PQVLVESPTV LEPGTKEX.
B_US_DH123  .......VGT PQILVESPAV LESGTKEX.
B_US_JRCSF  .......VGN PEILVESPTV LESGTKEX.
B_US_MNCG_  .......VGS PQILVESPTV LESGTKEX.
B_US_P896_  .......VGS PQILVESPAI LEPGTKEX.
B_US_RF_M1  .......VGS PQVLVESPAV LESGAKEX.
B_US_SF2_K  .......VGS PQILVESPAV LDSGTKEX.
B_US_WEAU1  .......VGS SQILLESPAV LEPGTKEX.
B_US_WR27_  .......VGD PQILGESPTV LGSGAKEX.
B_US_YU2_M  .......VGS PQILVESPPV LDSGTKEX.
BF1_BR_93B  .......VGS PQTSGESRAV LESGTKEX.
C_BR_92BR0  PQGNTERVGN PVFGRPCAVL ESRVKKEX.
```

```
C_BW_96BW0   SQGTTEGVGN PVSGKSCAIL GSRAKKEX.
C_BW_96BW1   SQGTPEGVGN PISGKSCAVL GARAKKEX.
C_BW_96BW1   SQGTTEGVGS PISGKSCAVL GSGTKKEX.
C_BW_96BW1   SQGTTEGVGS PVSGKSCAVL GSGTKKEX.
C_ET_ETH22   SQGTTEGVGN PISGKPCAVL GSGAKKEX.
C_IN_93IN1   .....L..GS PISGKSCAVL GSGAKKEX.
C_IN_93IN9   SQGTTERVGS PISGKSCAVL GSGAKKEX.
C_IN_93IN9   SQGTTEGVGS PISGKSCAVL GYRAKKEX.
C_IN_94IN1   SQGTTERVGS PISGKSCAVL GSGAKKEX.
C_IN_95IN2   SQGTTEGVGS PISGKSCTVL GSGAEKEX.
CRF01_AE_C   SQGTETGVGG PQISGESSVI LGSGTKNX.
CRF01_AE_C   STGTETEVGR PQISGESSVI LGSGTKNX.
CRF01_AE_C   SRGTETGVGR PQISGESSVI LGSGTENX.
CRF01_AE_T   SQGTETGVGR PQISGESSVI LGPGTKNX.
CRF01_AE_T   SQGTETGVGR PQISGESSVI LGSGTKNX.
CRF01_AE_T   SQGTETGVGR PQISGESSVI LGPGTKNX.
CRF01_AE_T   SQGTETGVGR PQISGESPVI LGPGTKNX.
CRF01_AE_T   SQGTETGVGR PQISGESSVI LGPGTKNX.
CRF01_AE_T   SQGTETGVGR PQISVESSGI LGPGTKNX.
CRF02_AG_F   PQGTETGVGS PPISGESSTI LGSGTKEX.
CRF02_AG_F   SQGTETGLGS PQISGESSDI LGAGTKEX.
CRF02_AG_G   .......VAD PQIPGESRAI LGSGTKEX.
CRF02_AG_N   SQGTETGVGS PQISVESYII LGSGTKEX.
CRF02_AG_S   .......VGS PQISVESSIV LGSGTKEX.
CRF02_AG_S   SQGTETGVGS SQTSVESSVI LGSGTKEX.
CRF03_AB_R   .......VGS PQILVESPTV LDSGTKEX.
CRF03_AB_R   .......VGS PQILVESPTV LDSGTKEX.
CRF04_cpx_   .......VGS PQVSVELPAV LGTGAKEX.
CRF04_cpx_   .......VGS PQVSVEPPAV LGTGAKEX.
CRF04_cpx_   .......VGN PQVPVEPPAV LGTGDKEX.
CRF05_DF_B   .......VGS PQISVEPPAI LESGTKEX.
CRF05_DF_B   .......VGS PQISVESPTV LESGAKEX.
CRF06_cpx_   .......VGN PQISGEPDML LGTGTTEX.
CRF06_cpx_   .......VGD PQIPGEPGVV LGTGTKEX.
CRF06_cpx_   .......VGD PQIPVEPGVL LGTGTKEX.
CRF06_cpx_   .......VGD PQIPGEPGVL LGTGTKEX.
CRF11_cpx_   .......VGD SQISGESDTV LGPRTEEX.
CRF11_cpx_   .......VGE SQIPGESSTV LGPRTEEX.
D_CD_84ZR0   .......VGS PQISVESPAI LESRTEEX.
D_CD_ELI_K   .......VGH PQISVESPTV LESGTEEQX
D_CD_NDK_M   .......VGS PQIPVEPPAV LESGTEEX.
D_UG_94UG1   .......VGS HQISVESPAV LDSGTKEX.
F1_BE_VI85   .......VGS SQISGESHAV LESGTKEX.
F1_BR_93BR   .......VGS SQISGESHTV LGSGTKEX.
F1_FI_FIN9   .......VGS PQISGEHHTV LESGTKEX.
F1_FR_MP41   .......VGN PQISMEPRTV LESGTKEX.
F2_CM_MP25   .......VGS PQISVESRAV LGSGTKEX.
F2KU_BE_VI   .......LGN PQIPVEPCAV LGSGTKEX.
G_BE_DRCBL   SQGTEIGVGS PQIFVESSVV LGSGTKEX.
G_NG_92NG0   PQGTETGVGR PQVLVEPPVV LGSGTKEX.
G_SE_SE616   PQGTETGVGR .SIFVESSVV LGQGTKEX.
H_BE_VI991   .......VGS PQTSGESPAV LGTGAKEX.
H_BE_VI997   .......KGG PQIPVESSTV LGTGTKEX.
H_CF_90CF0   .......EGS PQISLESSTI LGTGTKEX.
J_SE_SE702   .......VGD PQISGEPCMV LGAGTKEX.
J_SE_SE788   .......VGD PQISGEPCMV LGAGTKEX.
K_CD_EQTB1   .......LGS PQIPVEPDTV LGSGDKEX.
K_CM_MP535   .......LGS PQISVEPCTV LESGTKEX.
N_CM_YBF30   QQG.TATTET QNTLVGNTCI LGKRVKGX.
O_CM_ANT70   VDPR.AEDNC LQNLCSCNTI LATRIAEX.
O_CM_MVP51   VDPG.TKDNS LT.LWSCNAI LATRIEKX.
```

```
O_SN_MP129  VDSG.TEDNC LQTLHSCNTI LATRVAEX.
O_SN_MP130  VDPG.TEDNC LQNLHSCNTI LATRVAEX.
U_CD___83C  .......VGS TQIPGESCAV LGSGTKE..
```

Table 16. HIV Tat Alignment

GCG Multiple Sequence File.
Written by Omiga 1.1

Name: 00BW0762_1	Len:	108	Check:	5728	Weight:	1.00
Name: 00BW0768_2	Len:	108	Check:	4583	Weight:	1.00
Name: 00BW0874_2	Len:	108	Check:	5462	Weight:	1.00
Name: 00BW1471_2	Len:	108	Check:	4359	Weight:	1.00
Name: 00BW1616_2	Len:	108	Check:	5389	Weight:	1.00
Name: 00BW1686_8	Len:	108	Check:	6742	Weight:	1.00
Name: 00BW1759_3	Len:	108	Check:	6187	Weight:	1.00
Name: 00BW1773_2	Len:	108	Check:	5566	Weight:	1.00
Name: 00BW1783_5	Len:	108	Check:	6579	Weight:	1.00
Name: 00BW1795_6	Len:	108	Check:	6027	Weight:	1.00
Name: 00BW1811_3	Len:	108	Check:	4928	Weight:	1.00
Name: 00BW1859_5	Len:	108	Check:	6153	Weight:	1.00
Name: 00BW1880_2	Len:	108	Check:	6898	Weight:	1.00
Name: 00BW1921_1	Len:	108	Check:	6286	Weight:	1.00
Name: 00BW2036_1	Len:	108	Check:	4808	Weight:	1.00
Name: 00BW2063_6	Len:	108	Check:	7492	Weight:	1.00
Name: 00BW2087_2	Len:	108	Check:	4005	Weight:	1.00
Name: 00BW2127_2	Len:	108	Check:	6532	Weight:	1.00
Name: 00BW2276_7	Len:	108	Check:	7138	Weight:	1.00
Name: 00BW3819_3	Len:	108	Check:	4977	Weight:	1.00
Name: 00BW3842_8	Len:	108	Check:	5730	Weight:	1.00
Name: 00BW3871_3	Len:	108	Check:	7576	Weight:	1.00
Name: 00BW3876_9	Len:	108	Check:	4797	Weight:	1.00
Name: 00BW3886_8	Len:	108	Check:	7443	Weight:	1.00
Name: 00BW3891_6	Len:	108	Check:	5634	Weight:	1.00
Name: 00BW3970_2	Len:	108	Check:	5984	Weight:	1.00
Name: 00BW5031_1	Len:	108	Check:	8884	Weight:	1.00
Name: 96BW01B21	Len:	108	Check:	6237	Weight:	1.00
Name: 96BW0407	Len:	108	Check:	5097	Weight:	1.00
Name: 96BW0502	Len:	108	Check:	5303	Weight:	1.00
Name: 96BW06_J4	Len:	108	Check:	5679	Weight:	1.00
Name: 96BW11_06	Len:	108	Check:	7244	Weight:	1.00
Name: 96BW1210	Len:	108	Check:	5043	Weight:	1.00
Name: 96BW15B03	Len:	108	Check:	5056	Weight:	1.00
Name: 96BW16_26	Len:	108	Check:	5774	Weight:	1.00
Name: 96BW17A09	Len:	108	Check:	4288	Weight:	1.00
Name: 96BWMO1_5	Len:	108	Check:	6335	Weight:	1.00
Name: 96BWMO3_2	Len:	108	Check:	5085	Weight:	1.00
Name: 98BWMC12_2	Len:	108	Check:	5048	Weight:	1.00
Name: 98BWMC13_4	Len:	108	Check:	5714	Weight:	1.00
Name: 98BWMC14_a	Len:	108	Check:	6475	Weight:	1.00
Name: 98BWMO14_1	Len:	108	Check:	5085	Weight:	1.00
Name: 98BWMO18_d	Len:	108	Check:	4268	Weight:	1.00
Name: 98BWMO36_a	Len:	108	Check:	6665	Weight:	1.00
Name: 98BWMO37_d	Len:	108	Check:	5564	Weight:	1.00
Name: 99BW3932_1	Len:	108	Check:	5875	Weight:	1.00
Name: 99BW4642_4	Len:	108	Check:	6405	Weight:	1.00
Name: 99BW4745_8	Len:	108	Check:	4857	Weight:	1.00
Name: 99BW4754_7	Len:	108	Check:	5219	Weight:	1.00
Name: 99BWMC16_8	Len:	108	Check:	4870	Weight:	1.00
Name: A2_CD___97	Len:	108	Check:	4523	Weight:	1.00
Name: A2_CY___94	Len:	108	Check:	3933	Weight:	1.00
Name: A2D___97_9	Len:	108	Check:	4676	Weight:	1.00
Name: A2G_CD___9	Len:	108	Check:	2213	Weight:	1.00
Name: A_BY_97_97	Len:	108	Check:	4264	Weight:	1.00
Name: A_KE_93_Q2	Len:	108	Check:	3668	Weight:	1.00
Name: A_SE_93_SE	Len:	108	Check:	4159	Weight:	1.00

Name: A_SE_94_SE Len: 108 Check: 4323 Weight: 1.00
Name: A_SE_94_SE Len: 108 Check: 3099 Weight: 1.00
Name: A_SE_95_SE Len: 108 Check: 3717 Weight: 1.00
Name: A_SE_95_SE Len: 108 Check: 4178 Weight: 1.00
Name: A_SE_95_UG Len: 108 Check: 3954 Weight: 1.00
Name: A_UG_85_U4 Len: 108 Check: 3663 Weight: 1.00
Name: A_UG_92_92 Len: 108 Check: 4315 Weight: 1.00
Name: AC_IN_95_2 Len: 108 Check: 5100 Weight: 1.00
Name: AC_RW_92_9 Len: 108 Check: 4062 Weight: 1.00
Name: AC_SE_96_S Len: 108 Check: 6001 Weight: 1.00
Name: ACD_SE_95_ Len: 108 Check: 4767 Weight: 1.00
Name: ACG_BE___V Len: 108 Check: 5568 Weight: 1.00
Name: AD_SE_93_S Len: 108 Check: 4456 Weight: 1.00
Name: AD_SE_95_S Len: 108 Check: 4850 Weight: 1.00
Name: ADHK_NO_97 Len: 108 Check: 6557 Weight: 1.00
Name: ADK_CD_85_ Len: 108 Check: 4622 Weight: 1.00
Name: AG_BE___VI Len: 108 Check: 3720 Weight: 1.00
Name: AG_NG_92_9 Len: 108 Check: 4790 Weight: 1.00
Name: AGHU_GA___ Len: 108 Check: 4548 Weight: 1.00
Name: AGU_CD_76_ Len: 108 Check: 7456 Weight: 1.00
Name: AJ_BW_98_B Len: 108 Check: 4086 Weight: 1.00
Name: B_AU___VH_ Len: 108 Check: 5824 Weight: 1.00
Name: B_CN___RL4 Len: 108 Check: 3342 Weight: 1.00
Name: B_DE_86_D3 Len: 108 Check: 4538 Weight: 1.00
Name: B_DE_86_HA Len: 108 Check: 5673 Weight: 1.00
Name: B_FR_83_HX Len: 108 Check: 2953 Weight: 1.00
Name: B_GA___OYI Len: 108 Check: 5056 Weight: 1.00
Name: B_GB___CAM Len: 108 Check: 4131 Weight: 1.00
Name: B_GB___GB8 Len: 108 Check: 7783 Weight: 1.00
Name: B_GB_59_MA Len: 108 Check: 5562 Weight: 1.00
Name: B_KR___WK_ Len: 108 Check: 6702 Weight: 1.00
Name: B_NL_86_32 Len: 108 Check: 5027 Weight: 1.00
Name: B_TW___TWC Len: 108 Check: 6611 Weight: 1.00
Name: B_US___DH1 Len: 108 Check: 5131 Weight: 1.00
Name: B_US___P89 Len: 108 Check: 5087 Weight: 1.00
Name: B_US_83_RF Len: 108 Check: 7745 Weight: 1.00
Name: B_US_83_SF Len: 108 Check: 4303 Weight: 1.00
Name: B_US_84_MN Len: 108 Check: 4888 Weight: 1.00
Name: B_US_86_JR Len: 108 Check: 5959 Weight: 1.00
Name: B_US_86_YU Len: 108 Check: 6104 Weight: 1.00
Name: B_US_87_BC Len: 108 Check: 5602 Weight: 1.00
Name: B_US_88_WR Len: 108 Check: 5031 Weight: 1.00
Name: B_US_90_WE Len: 108 Check: 5035 Weight: 1.00
Name: BF1_BR_93_ Len: 108 Check: 3381 Weight: 1.00
Name: C_BR_92_92 Len: 108 Check: 6035 Weight: 1.00
Name: C_BW_96_96 Len: 108 Check: 5570 Weight: 1.00
Name: C_BW_96_96 Len: 108 Check: 7348 Weight: 1.00
Name: C_BW_96_96 Len: 108 Check: 5043 Weight: 1.00
Name: C_BW_96_96 Len: 108 Check: 5056 Weight: 1.00
Name: C_ET_86_ET Len: 108 Check: 4199 Weight: 1.00
Name: C_IN_93_93 Len: 108 Check: 5957 Weight: 1.00
Name: C_IN_93_93 Len: 108 Check: 4573 Weight: 1.00
Name: C_IN_93_93 Len: 108 Check: 5361 Weight: 1.00
Name: C_IN_94_94 Len: 108 Check: 5479 Weight: 1.00
Name: C_IN_95_95 Len: 108 Check: 5697 Weight: 1.00
Name: CRF01_AE_C Len: 108 Check: 2633 Weight: 1.00
Name: CRF01_AE_C Len: 108 Check: 4093 Weight: 1.00
Name: CRF01_AE_C Len: 108 Check: 2957 Weight: 1.00
Name: CRF01_AE_T Len: 108 Check: 2478 Weight: 1.00
Name: CRF01_AE_T Len: 108 Check: 1857 Weight: 1.00
Name: CRF01_AE_T Len: 108 Check: 1204 Weight: 1.00

```
Name: CRF01_AE_T      Len:   108  Check: 2024  Weight:   1.00
Name: CRF01_AE_T      Len:   108  Check: 913   Weight:   1.00
Name: CRF01_AE_T      Len:   108  Check: 1446  Weight:   1.00
Name: CRF02_AG_F      Len:   108  Check: 4840  Weight:   1.00
Name: CRF02_AG_F      Len:   108  Check: 6283  Weight:   1.00
Name: CRF02_AG_G      Len:   108  Check: 4683  Weight:   1.00
Name: CRF02_AG_N      Len:   108  Check: 3989  Weight:   1.00
Name: CRF02_AG_S      Len:   108  Check: 3401  Weight:   1.00
Name: CRF02_AG_S      Len:   108  Check: 4884  Weight:   1.00
Name: CRF03_AB_R      Len:   108  Check: 2749  Weight:   1.00
Name: CRF03_AB_R      Len:   108  Check: 3199  Weight:   1.00
Name: CRF04_cpx_      Len:   108  Check: 2779  Weight:   1.00
Name: CRF04_cpx_      Len:   108  Check: 2140  Weight:   1.00
Name: CRF04_cpx_      Len:   108  Check: 4857  Weight:   1.00
Name: CRF05_DF_B      Len:   108  Check: 6229  Weight:   1.00
Name: CRF05_DF_B      Len:   108  Check: 4828  Weight:   1.00
Name: CRF06_cpx_      Len:   108  Check: 4652  Weight:   1.00
Name: CRF06_cpx_      Len:   108  Check: 5232  Weight:   1.00
Name: CRF06_cpx_      Len:   108  Check: 4871  Weight:   1.00
Name: CRF06_cpx_      Len:   108  Check: 5747  Weight:   1.00
Name: CRF11_cpx_      Len:   108  Check: 4101  Weight:   1.00
Name: CRF11_cpx_      Len:   108  Check: 3288  Weight:   1.00
Name: D_CD_83_EL      Len:   108  Check: 3640  Weight:   1.00
Name: D_CD_83_ND      Len:   108  Check: 4658  Weight:   1.00
Name: D_CD_84_84      Len:   108  Check: 6447  Weight:   1.00
Name: D_UG_94_94      Len:   108  Check: 6559  Weight:   1.00
Name: F1_BE_93_V      Len:   108  Check: 4812  Weight:   1.00
Name: F1_BR_93_9      Len:   108  Check: 4376  Weight:   1.00
Name: F1_FI_93_F      Len:   108  Check: 4918  Weight:   1.00
Name: F1_FR_96_M      Len:   108  Check: 3934  Weight:   1.00
Name: F2_CM_95_M      Len:   108  Check: 5318  Weight:   1.00
Name: F2KU_BE_94      Len:   108  Check: 4644  Weight:   1.00
Name: G_BE_96_DR      Len:   108  Check: 4391  Weight:   1.00
Name: G_NG_92_92      Len:   108  Check: 5855  Weight:   1.00
Name: G_SE_93_SE      Len:   108  Check: 2420  Weight:   1.00
Name: H_BE___VI9      Len:   108  Check: 5780  Weight:   1.00
Name: H_BE___VI9      Len:   108  Check: 2472  Weight:   1.00
Name: H_CF_90_90      Len:   108  Check: 3026  Weight:   1.00
Name: J_SE_93_SE      Len:   108  Check: 4542  Weight:   1.00
Name: J_SE_94_SE      Len:   108  Check: 4057  Weight:   1.00
Name: K_CD_97_EQ      Len:   108  Check: 2222  Weight:   1.00
Name: K_CM_96_MP      Len:   108  Check: 4681  Weight:   1.00
Name: N_CM_95_YB      Len:   108  Check: 3679  Weight:   1.00
Name: O_CM___ANT      Len:   108  Check: 1276  Weight:   1.00
Name: O_CM_91_MV      Len:   108  Check: 814   Weight:   1.00
Name: O_SN___99S      Len:   108  Check: 9132  Weight:   1.00
Name: O_SN___99S      Len:   108  Check: 2342  Weight:   1.00
Name: U_CD___83C      Len:   108  Check: 4773  Weight:   1.00

                1                                                  50
00BW0762_1  ...MEQVDPN LEPWNHPGSQ PKTPCNKCFC KSCSYHCLVC FQTKGLGISY
00BW0768_2  ...MDPVDPN LEPWNHPGSQ PKTACNNCYC KYCSYHCLVC FQKKGLGISY
00BW0874_2  ...MEPVDPN LEPWNHPGSQ PKTACNSCYC KHCCYHCRVC FLTKGLGISY
00BW1471_2  ...MDPVDPK LEPWNHPGSQ PKTACNTCYC KKCCYHCQVC FLNKGLGISY
00BW1616_2  ...MEPIDPN LEPWNHPGSQ PSTACNKCYC KQCCYHCPAC FLTKGLGISY
00BW1686_8  ...MEPVDPK LEPWNHPGSQ PTTPCTTCFC KVCSYHCLVC FQTKGLGIYY
00BW1759_3  ...MEPIDPN LEPWNHPGSQ PKTACNKCYC KSCSYHCLVC FQKKGLGISY
00BW1773_2  ...MEPVDPN LEPWNHPGSQ PKTPCTKCYC KLCSYHCLVC FQTKGLSISY
00BW1783_5  ...MEPVDPN LEPWNHPGSQ PKTPCTKCYC KKCSYHCLVC FQTKGLGISY
00BW1795_6  ...MEPVDPS LEPWNHPGSQ PKTPCNHCYC KHCSYHCLVC FQRKGLGISY
00BW1811_3  ...MEPVDPN LEPWKHPGSQ PKTACNTCYC KRCSYHCPVC FLTKGLGISY
```

```
00BW1859_5  ...MEPVDPN LEPWNHPGSR PKTACNKCYC KRCSYHCLVC FQTKGLGISY
00BW1880_2  ...MEPVDPN LEPWNHPGSQ PQTPCNKCYC KRCSYHCLVC FQTKGLGISY
00BW1921_1  ...MEPIDPK LEPWNHPGSQ PNTPCTPCYC KKCSYHCLVC FQTKGLGISY
00BW2036_1  ...MEPVDPK LEPWNHPGSQ PKTPCNKCYC KKCSYHCLVC FQTKGLGIYY
00BW2063_6  ...MEIVDPN LDPWDHPGSQ PSTPCNKCYC KYCSYHCLVC FQTKGLGISY
00BW2087_2  ...MEPIDPN LEPWNHPGSQ PKTACNKCFC KRCCYHCPAC FLTKGLGISY
00BW2127_2  ...MEPVDPN LEPWNHPGSQ PKTACSPCYC KNCSYHCLVC FQTKGLGISY
00BW2276_7  ...MEPVDPN LEPWKHPGSQ PKTACTSCYC KRCSYHCLVC FQTKGLGISY
00BW3819_3  ...MEPVDPN LEPWNHPGSQ PKTACNQCYC KHCCYHCLVC FQTKGLGISY
00BW3842_8  ...MEPVDPK LEPWNHPGSQ PRTPCNKCFC KVCSYHCQSC FLTKGLGISY
00BW3871_3  ...MDPVDPN LEPWNHPGSQ PKTPCNNCYC KKCSYHCVVC FQTKGLGIYY
00BW3876_9  ...MELIDPK LEPWNHPGSQ PAAPCNNCYC KHCSYHCLVC FQTKGLGISY
00BW3886_8  ...MEPIDPN LEPWNHPGSQ PKTPCNKCYC KHCSYHCLVC FQNKGLGISY
00BW3891_6  ...MEPVDPN LEPWNHPGSQ PNTACNKCYC KQCSYHCLVC FQKKGLGISY
00BW3970_2  ...MEPVDPN LEPWNHPGSQ PKTACNSCYC KKCSYHCQVC FLTKGLGISY
00BW5031_1  ...MEPVDPN LEPWNHPGSQ PKTPCNKCYC KQCSYHCPVC FLTKGLGIYY
 96BW01B21  ...MEPVDPN LEPWNHPGSQ PKTACNNCYC KHCSYHCLVC FQTKGLGIYY
  96BW0407  ...MEPIDPN LEPWNHPGSQ PKTACNKCYC KHCSYHCLVC FQTKGLGISY
  96BW0502  ...MEPVDPN LEPWKHPGSQ PRTACTKCYC KYCCYHCLVC FQTKGLGISY
 96BW06_J4  ...MEPVDPK LEPWNHPGGQ PKTPCNTCYC KKCSYHCLVC FQKKGLGIYY
 96BW11_06  ...MEQVDPN LEPWNHPGSQ PKTPCNNCYC KYCSYHCLVC FQTKGLGISY
  96BW1210  ...MEPVDPN LEPWNHPGSQ PKIACNKCYC KACCYHCLVC FQTKGLGISY
 96BW15B03  ...MEPVDPK LEPWNHPGSQ PKTPCTKCFC KGCSYHCLVC FQTKGLGISY
 96BW16_26  ...MEPVDPK LEPWNHPGSQ PKTACNKCYC KRCSYHCLVC FQTKGLGISY
 96BW17A09  ...MDPVDPS LEPWNHPGSQ PKTACNTCYC KKCCYHCQVC FLNKGLGISY
 96BWMO1_5  ...MEPVDPN LEPWKHPGSQ PETPCNKCYC KSCSYHCLVC FQKKGLGISY
 96BWMO3_2  ...MEPVDPN LEPWNHPGSQ PKTACNNCCC KRCSYHCLVC FQKKGLGISY
98BWMC12_2  ...MEPVDPN LEPWNHPGSQ PKTACNNCYC KRCCYHCQRC FLTKGLGISY
98BWMC13_4  ...MEPVDPN LEPWNHPGSQ PKTSCNNCYC KKCSYHCLVC FQKKGLGISY
98BWMC14_a  ...MDPVDPK LEPWNHPGSQ PSTPCNTCFC KKCSYHCLVC FQKKGLGIYY
98BWMO14_1  ...MEPIDPN LEPWNHPGSQ PKTACNKCHC KRCCYHCTVC FLTKGLGISY
98BWMO18_d  ...MEPIDPN LEPWNHPGSQ PKTDCNKCYC KRCCYHCLVC FQKKGLGISH
98BWMO36_a  ...MEPVDPN LEPWNHPGSQ PRTACNTCYC KKCSYHCLVC FQKKGLGISY
98BWMO37_d  ...MEPVDPN LEPWNHPGSQ PKTACNKCYC KHCSYHCLVC FQTKGLGISY
99BW3932_1  ...MEPVDPN LEPWNHPGSQ PKTACNKCYC KSCSYHCPVC FLSKGLGISY
99BW4642_4  ...MEPVDPN LEPWNHPGSQ PKTPCTTCFC KRCSYHCLVC FQTKGLGIYY
99BW4745_8  ...MEPVDPN LEPWNHPGSQ PSTACNKCYC KHCSYHCLVC FQTKGLAISY
99BW4754_7  ...MEPVDPN LEPWNHPGSQ PETPCNKCFC KACSYHCLVC FQKKGLGISY
99BWMC16_8  ...MDPVDPN LEPWNHPGSQ PKTACNRCYC KKCSYHCLVC FHTKGLGISY
A2_CD___97  ...MEPVDPN LEPWNHPGSQ PKTACNNCYC KKCCYHCQRC FLNKGLGISY
A2_CY___94  ...MEPVDPK LEPWNHPGSQ PKTACTKCYC KRCCYHCQLC FINKGLGISY
A2D___97_9  ...MDPVDPN LEPWNHPGSQ PRTACNKCYC KRCCYHCPVC FLNKGLGISY
A2G_CD___9  ...MDPVDPN LEPWNHPGSQ PKTACNKCYC KACCWHCQVC FLNKGLGISY
A_BY_97_97  ...XDPVDPN LEPXNHXXSQ PKTXCRNCYC KKCCWHCQIC FLKKGLGISY
A_KE_93_Q2  ...MDPVDPN LEPWNHPGSQ PTTACNKCYC KKCCYHCQVC FLNKGLGISY
A_SE_93_SE  ...MDPVDPN LEPWNHPGSQ PSTACNKCYC KICCYHCPVC FLNKGLGISY
A_SE_94_SE  ...MDPVDPN LEPWNHPGSQ PTTPCSKCYC KKCCYHCIVC FQNKGLGISY
A_SE_94_SE  ...MDPVDPN LEPWNHPGSQ PTTACNKCYC KKCCYHCQAC FLNKGLGISY
A_SE_95_SE  ...MEPVDPN IEPWNHPGSQ PTTPCTKCYC KKCCYHCPAC FLNKGLGISY
A_SE_95_SE  ...MDPVDPN LEPWNHPGSQ PATACSACYC KKCCYHCPVC FLNKGLGISY
A_SE_95_UG  ...MDPVDPN LEPWNHPGSQ PTTPCNKYFC KRCCYHCLVC FQHKGLGISY
A_UG_85_U4  ...MEPVDPN LEPWKHPGSQ PTTACSNCYC KVCCWHCQLC FLKKGLGISY
A_UG_92_92  ...MDPVDPS LEPWNHPGSQ PKTPCNKCYC KVCCYHCQCC FLNKGLGISY
AC_IN_95_2  ...MEPVDPN LEPWNHPGSQ PRTACNNCYC KRCSYHCLVC FQKKGLGISY
AC_RW_92_9  ...MEPVDPK LEPWNHPGSQ PKTACNNCYC KHCSYHCLVC FQAKGLGISY
AC_SE_96_S  ...MEPIDPN LEPWNHPGSQ PKTACNRCYC KRCSYHCLVC FQTKGLGISY
ACD_SE_95_  ...MDLVDPN LEPWNHPGSQ PTTPCSKCYC KKCCYHCQHC FITKGLGISY
ACG_BE___V  ...MEPVDPK LEPWNHPGSQ PKTACNKCFC KKCSYHCLVC FQTKGLGISY
AD_SE_93_S  ...MDPVDPS LEPWNHPGSQ PTTPCTKCYC KKCCYHCQVC FITKGLGISY
AD_SE_95_S  ...MDPVDPN LEPWNHPGSQ PTTPCSKCYC KKCCYHCQSC FLNKGLGISY
ADHK_NO_97  ...MDPVDHN LEPWNQPGSQ PNTPCNNCYC KQCCYHCQLC FLRKGLGISY
```

```
ADK_CD_85_   ...MDPVDPN LEPWNHPGSQ PRTPCNKCYC KKCCYHCQMC FITKGLGISY
AG_BE___VI   ...MDPVDPN LEPWNHPGSQ PKTACNKCYC KKCCWHCQVC FLNKGLGISY
AG_NG_92_9   ...MELVDPS LEPWNHPGSQ PTTACNKCYC KICCWHCQLC FLNKGLGISY
AGHU_GA___   ...MDPVDPN IEPWNQPGSQ PKTACNQCYC KRCCYHCQHC FLKKGLRISY
AGU_CD_76_   ...MEPVDPN LEPWNQPGSQ PQTACNNCYC KKCCFHCQVC FSSKGLGISY
AJ_BW_98_B   ...MEPVDPN IEPWNQPGSQ PKTACTPCYC KKCCFHCQVC FLKKGLGISY
B_AU___VH_   ...MDPVDPR LEPWKHPGSQ PKTACTNCYC KQCCFHCQVC FIRKGLGISY
B_CN___RL4   ...MEPVDPR LEPWKHPGSQ PKTACNNCYC KKCCLHCQVC FMKKGLGIFY
B_DE_86_D3   ...MEPVDPR LEPWEHPGSQ PKTACTNCYC KKCCFHCQVC FTKKGLGISY
B_DE_86_HA   ...MEPVDPR LEPWKHPGSQ PRTPCTNCYC KKCCLHCQVC FTRKGLSISY
B_FR_83_HX   ...MEPVDPR LEPWKHPGSQ PKTACTNCYC KKCCFHCQVC FITKALGISY
B_GA___OYI   ...MEPVDPR LEPWKHPGSQ PKTASNNCYC KRCCLHCQVC FTKKGLGISY
B_GB___CAM   ...MEPVDPR LEPWKHPGSQ PRTACTNCYC KQCCFHCQVC FITKGLGISY
B_GB___GB8   ...MEPVDPR IEPWKHPGSQ PRTACTNCYC KKCCFHCQVC FLRKGLGISY
B_GB_59_MA   ...MEPVDPR LEPWNHPGSQ PKTACTKCYC KKCCLHCQVC FITKGLGISY
B_KR___WK_   ...MEPVDPR LEPWKHPGSQ PKTPCTKCYC KKCCLHCQVC FMTKGLGISY
B_NL_86_32   ...MEPVDPK LEPWKHPGSQ PKTACNNCYC KKCCFHCQVC FTKKGLGISY
B_TW___TWC   ...MDPVDPR LEPWKHPGSQ PKTACTNCYC KKCCLHCQVC FMTKALGISN
B_US___DH1   ...MEPVDPR LEPWKHPGSQ PKTACTNCYC KKCCLHCQVC FITKGLGISY
B_US___P89   ...MEPVNPS LEPWKHPGSQ PKTACTNCYC KKCCFHCQAC FITKGLGISY
B_US_83_RF   ...MEPVDPR LEPWKHPGSQ PKTACNNCYC KKCCYHCQVC FLTKGLGISY
B_US_83_SF   ...MEPVDPN LEPWKHPGSQ PRTACNNCYC KKCCFHCYAC FTRKGLGISY
B_US_84_MN   ...MEPVDPR LEPWKHPGSQ PKTACTTCYC KKCCFHCQVC FTKKALGISY
B_US_86_JR   ...MEPVDPS LEPWKHPGSQ PKTACTNCYC KKCCLHCQVC FTTKGLGISY
B_US_86_YU   ...MEPVDPN LEPWKHPGSQ PRTACNNCYC KKCCFHCQVC FTKKGLGISY
B_US_87_BC   ...MEPVDPR LEPWKHPGSQ PKTACTKCYC KKCCLHCQVC FMTKGLGISY
B_US_88_WR   ...MDPVDPR LEPWKHPGSQ PKTACTNCYC KKCCFHCQVC FITKGLGISY
B_US_90_WE   ...MEPVDHR LEPWKHPGSQ PKTACTNCYC KRCCFHCQVC FMTKGLGISY
BF1_BR_93_   ...MEPVDPR LEPWKHPGSR PQTACNSCYC KKCCFHCQVC FTTKGLGISY
C_BR_92_92   ...MEPVDPN LEPWNHPGSQ PKTACNNCYC KRCSYHCLVC FQTKGLGISY
C_BW_96_96   ...MEPIDPN LEPWNHPGSQ PKTACNKCFC KHCSYHCLVC FQTKGLSISY
C_BW_96_96   ...MEQVDPN LEPWNRPGSQ PKTPCNNCYC KYCSYHCLVC FQTKGLGISY
C_BW_96_96   ...MEPVDPN LEPWNHPGSQ PKIACNKCYC KACCYHCLVC FQTKGLGISY
C_BW_96_96   ...MEPVDPK LEPWNHPGSQ PKTPCTKCFC KGCSYHCLVC FQTKGLGISY
C_ET_86_ET   ...MEPVDPN LEPWNHPGSQ PKTACNQCYC KKCSYHCLVC FLTG.LGISY
C_IN_93_93   ...MEPVDPN LEPWNHPGSQ PKTACNQCYC KRCSYHCLVC FQKKGLGISY
C_IN_93_93   ...MEPVDPN LEPWNHPGSQ PKTACNNCYC KHCSYHCLVC FQKKGLGISY
C_IN_93_93   ...MEPVDPN LEPWNHPGSQ PRTACNNCFC KSCSYHCLVC FQKKGLGISY
C_IN_94_94   ...MEPIDPN LEPWNHPGSQ PKTACNNCYC KHCSYHCLVC FQRKGLGISY
C_IN_95_95   ...MEPVDPN LEPWNHPGSQ PKTACNNCYC KKCSYHCLVC FQKKGLGISY
CRF01_AE_C   ...MEPVDPS LEPWNHPGSQ PTTACSKCYC KKCCWHCQLC FLKKGLGISY
CRF01_AE_C   ...MEPVDPN LEPWKHPGSQ PTTACSKCYC KMCCWHCQLC FLKKGLGISY
CRF01_AE_C   ...MELVDPK LEPWNHPGSQ PTTACNKCYC KICCWHCQLC FLKKGLGISY
CRF01_AE_T   ...MELVDPN LEPWNHPGSQ PTTACSKCYC KKCCWHCQLC FLKKGLGISY
CRF01_AE_T   ...MELVDPN LEPWKHPGSQ PSTACSKCYC KICCWHCQLC FLKKGLGISY
CRF01_AE_T   ...MEPVDPN LEPWNHPGSQ PTTACSKCYC KICCWHCQLC FLKKGLGISH
CRF01_AE_T   ...MELVDPN LEPWNHPGSQ PTTACSKCYC KKCCWHCQLC FLKKGLGISH
CRF01_AE_T   ...MELVDPN LEPWNHPGSQ PTTACSKCYC KKCCWHCQLC FLKKGLGISH
CRF01_AE_T   ...MELVDPN LEPWNHPGSQ PTTACSKCYC KKCCWHCQLC FLKKGLGISH
CRF02_AG_F   ...MEPVDPS LEPWNHPGSQ PTTACSKCYC KICCWHCQLC FLNKGLGISY
CRF02_AG_F   ...MEPVDPS LEPWNHPGSQ PTTACSNCYC KICCWHCQLC FLNKGLGISY
CRF02_AG_G   ...MELIDPS LEPWNHPGSQ PTTACSKCYC KICCWHCQLC FLNKGLGISY
CRF02_AG_N   ...MELVDPS LEPWNHPGSQ PTTACSKCYC KMCC.HCQLC FLNKGLGISY
CRF02_AG_S   ...MELVDPS LEPWNHPGSQ PTTACSKCYC KICCWHCQLC FLNKGLGISY
CRF02_AG_S   ...MELVDPK LEPWNHPGSQ PTTACSKCYC KLCCWHCQLC FLNKGLGISY
CRF03_AB_R   ...MEPVDPR LEPWKHPGSQ PKTACTNCYC KKCCLHCQVC FMKKGLGISY
CRF03_AB_R   ...MEPVDPR LEPWKHPGSQ PKTACTNCYC KKCCLHCQVC FMKKGLGISY
CRF04_cpx_   ...MEPVDPD LEPWNHPGSQ PTTDCNKCFC KKCCWHCQVC FLKKGLGISY
CRF04_cpx_   MEAVDLVYPN LEPWNHPGSQ PTTACNKCFC KKCCWHCQAC FLKKGLGISY
CRF04_cpx_   ...MEPVDPN LEPWNHPGSQ PTTACTKCFC KVCCYHCQVC FLRKGLGISY
CRF05_DF_B   ...MDPVDPN LEPWNHPGSQ PRTACNQCYC KMCCYHCQNC FITKGLGISY
```

```
CRF05_DF_B  ...MDPVDPN LEPWNHPGSQ PRTACNQCHC KKCCYHCQVC FITKGLGISY
CRF06_cpx_  ...MEPVDPK IEPWNQPGSR PKTACTKCYC KKCCYHCPVC FLNKGLGISY
CRF06_cpx_  ...MEPVDPN IEPWNQPGSR PKTACTPCYC KKCSYHCQLC FLNKGLGISY
CRF06_cpx_  ...MEPVDPK IEPWNQPGSR PKTACTSCYC KRCCYHCPLC FLNKGLGISY
CRF06_cpx_  ...MEPVDPN IEPWNQPGSR PKTACTSCYC KQCCYHCPVC FLNKGLGISY
CRF11_cpx_  ...MEPVDPN IEPWNQPGSR PKTACNQCYC KRCCYHCPHC FLKKGLGIYY
CRF11_cpx_  ...MEPGDIN IDPWNQPGSQ PKTACNQCYC KRCCYHCQHC FLKKGLGISY
D_CD_83_EL  ...MDPVDPN LEPWNHPGSQ PRTPCNKCHC KKCCYHCPVC FLNKGLGISY
D_CD_83_ND  ...MDPVDPN LESWNHPGSQ PRTACNKCHC KKCCYHCQVC FITKGLGISY
D_CD_84_84  ...MDPVDPN IDPWNHPGSQ PRTACNNCYC KKCCYHCQVC FITKGLGISY
D_UG_94_94  ...MEPVDPN LEPWNHPGSQ PRTPCNKCYC KRCCYHCYVC FVTKGLGISY
F1_BE_93_V  ...MEPVDPS LDPWNHPGSQ PTTPCTKCYC KRCCFHCQWC FTTKGLGISY
F1_BR_93_9  ...MELVDPN LDPWNHPGSQ PTTPCTRCYC KWCCFHCYWC FTTKGLGISY
F1_FI_93_F  ...MELVDPN LDPWNHPGSQ PPTPCNKCYC KRCCFHCYWC FATKGLGISY
F1_FR_96_M  ...MELVDPN LDPWNQPGSQ PTTPCTKCYC KKCCFHCYCC FATKGLGISY
F2_CM_95_M  ...MEVVDPN LDPWKHPGSQ PETPCNKCYC KKCCFHCQLC FTRKGLGISY
F2KU_BE_94  ...MEQVDPN LDPWNHPGSQ PKTACNNCYC KKCCFHCQVC FTTKGLGISY
G_BE_96_DR  ...MDPVDPK LEPWKHPGSQ PKTPCNNCYC KS.VAALQVC FLNKGLGISY
G_NG_92_92  ...MDPVDPK LEPWNHPGSQ PTTPCNKCYC KVCCWHCQVC FLNKGLGISY
G_SE_93_SE  ...MDPVDPN LEPWNHPGSQ PKTPCNKCFC KVCCWHCQVC FLNKGLGISY
H_BE___VI9  ...MDPVDPN QEPWNHPGSQ PRTACNNCYC KKCCYHCQLC FLKKGLGIYY
H_BE___VI9  ...MDPVDPN QEPWNHPGNQ PKTPCNNCYC KQCSFHCQLC FLKKGLGISY
H_CF_90_90  ...MDPVDPK LEPWNHPGSQ PQTACNNCYC KKCCYHCQMC FLKKGLGISY
J_SE_93_SE  ...MEPVDPN REPWNHPGSQ PKTACTNCYC KKCCYHCQVC FLQKGLGISY
J_SE_94_SE  ...MEPVDPN REPWNHPGSQ PKTACTNCYC KKCCYHCQVC FLQKGLGISY
K_CD_97_EQ  ...MEPVDPN IEPWNQPGSQ PKTACNQCYC KKCCYHCQLC FLQKGLGICY
K_CM_96_MP  ...MDPVDPN IEPWNQPGSQ PKTACNQCYC KRCCYHCQIC FLKKGLGISN
N_CM_95_YB  ...MEPVDPR LEPWNHPGSQ PKTACNNCYC KRCCYHCLYC FTKKGLGISY
O_CM___ANT  ...MDPVDPE VPPWHHPGSQ PQIPCNNCYC KRCCYHCYVC FVRKGLGISY
O_CM_91_MV  ...MDPVDPE MPPWHHPGSK PQTPCNNCYC KRCCYHCYVC FTKKGLGISH
O_SN___99S  ...MDPVDPE MPPWHHPGSQ PQTPCNKCYC KACCYHCYVC FASKGLGISY
O_SN___99S  ...MDPVDPE MPPWHHPGSQ PQTPCNKCYC KACCYHCYVC FASKGLGLSY
U_CD___83C  ...MDPVDPR LEPWNHPGSQ PKTACNSCYC KRCCLHCQVC FMTKGLGISY

            51                                                  100
00BW0762_1  GRKKRRQRR. .HAPPGSKDH QNPISKQPLP .QTRGDPTGS EESKKKVESK
00BW0768_2  GRKKRGQRR. .NAPPSSEDH QDLISKQPLS .QTRGDPTGS EESKKKAESK
00BW0874_2  GRKKRRQRR. .SAPPSGEDH QSSLPKQPLP .QARGNPTGS EESKKKVESK
00BW1471_2  GRKKRRQRR. .SAPPSSKDH QDPLSKQPLS .QTRGDPTGS KESKKKVESK
00BW1616_2  GRKKRRQRR. .STPPSSKDH QNPVSKQPLP .RTQGNSTGS EESKKKVESK
00BW1686_8  GRKKRRQRR. .SAPPSNKEH QDPIPKQPLP .RTPGDPTGP EESKKKVESK
00BW1759_3  GRKKRRQRR. .SGPPSNKDH QNLVSKQPLS .QTQRDPTGS EESKKKVESK
00BW1773_2  GRKKRRQRR. .STPPSSEDH QDLISKQPLP .RTQGDPTGS EESKKKVESK
00BW1783_5  GRKKRRQRR. .SGPPSSEDH QNPISKQPLS .QTRGDPTGS KESKKKVESK
00BW1795_6  GRKKRRRRR. .SAPPSSEDH QNPISKQPLS .QTRGDPTGS EESKKKVESK
00BW1811_3  GRKKRRQRR. .RAPPSSEDH QNPIPKQPLS .RTQGDPTGS EESKKKVESK
00BW1859_5  GRKKRGQRR. .SAPSSSEDH QDNISKQPLP .RTQGDPTGS EESKKKVESK
00BW1880_2  GRKKRRQRR. .SAPPSSEDH QNPISKQPFP .QTRGDSTGS EESKKKVESK
00BW1921_1  GRKKRRQRR. .SAPPSSKDH QNPLSKQPLP .RTPGNSTGS EESKKKVESK
00BW2036_1  GRKKRGQRR. .STPPSNKAH QNLISKQPLP .RTQGDSTGS EESKKKVESK
00BW2063_6  GRKKWRQRRG RSAPSSSKDH QDLISKQPLP .QTRGDPTSS KESKKKVESK
00BW2087_2  GRKKRRQRR. .SAPPGSEDH QNPISKQPLP .QTRGNPTGS EESKKKVESK
00BW2127_2  GRKKRRQRR. .SAPPSSEDH QDPISKQPLP .QTRGDPTGS EEPKKKVESK
00BW2276_7  GRKKRRQRR. .SAPPSSEDH QNHISKQPLP .QTRGDQTGS EESKKKVESK
00BW3819_3  GRKKRRQRR. .RSPQSSEAH QDPISKQPLP .QARGDPTGS EESKKKVESK
00BW3842_8  GRKKRGQRR. .SAPSSSEDH QNPIPKQPLP .RTQGNSTGS EESKKKVESK
00BW3871_3  GRKKRRQRR. .SAPPSNKDH QNPISKQPLP .RTQGNSTGS EESKKKVESK
00BW3876_9  GRKKRRQRR. .STPPRSEDH QNLILKQPLP .RTQGNPTGP EESKKKVESK
00BW3886_8  GRKKRRQRR. .SAPPSSEDH QNLISKQPLP .RTPGNSTGS EESKKKVGSK
00BW3891_6  GRKKRRQRR. .SAPPSSEDH QNPISKQPLS .QTRGDSTGS EESKKKVESK
00BW3970_2  GRKKRRQRR. .SAPPSSEDH QNLISKRPLP .QTRGNQTGS EESKKKVESK
```

```
00BW5031_1  GRKKRRQRR. .SAPSSSENH QNPVSKQPLP .QARGNSTGS EESKKKVESK
 96BW01B21  GRKKRRQRR. .SAPPSSEDH QNLISKQPLP .RTQGDPTGS EESKKKVESK
  96BW0407  GRKKRRQRR. .STPPSSEDH QNPVSKQPLS .QTRGDPTGS EESKKKVESK
  96BW0502  GRKKRRQRR. .STPPSSESH QNLISEQPLP .RTQGNPTGS EESKKKVESK
 96BW06_J4  GRKKRRQRR. .SAPSSSEDH QDPISKQPFP .RTQGESTGS EESKKEVESK
 96BW11_06  GRKKRRQRR. .SAPPSSKDH QNPVSKQPLP .QTRGDPTGS KESKRKVENK
  96BW1210  GRKKRRQRR. .SAPPSSEDH QDLVPKQPLS .QARGNPTSS KESKKKVESK
 96BW15B03  GRKKRGQRR. .SAPPRSEDH QNLISKQPLP .RTQGDSTGS EESKKKVESK
 96BW16_26  GRKKRRQRR. .SAPPSSEDH QNPISKQPLP .QTQGDPTGS EESKKKVESK
 96BW17A09  GRKKRRQRR. .SAPPSNGDH QNPISKQPLP .QTRGDPTGS KESKKEVESK
 96BWMO1_5  GRKKRRQRR. .SAPPSSEDH QDLVSKQPLP .QTRGDPTGS KESKKKVESK
 96BWMO3_2  GRKKRGQRR. .SAPPSSKDH QNPVSKQPLP .QTRGNSTGS EESKKKVESK
98BWMC12_2  GRKKRRQRR. .SAPPSCKDH QNPLPEQPST .QTRGDSTGS EESKKKVESK
98BWMC13_4  GRKKRRQRR. .SAPASSEDH QDLISKQLLP .QTRGDPTGS KESKKKVESK
98BWMC14_a  GRKKRRQRR. .STPPSSKDH QNPVSKQPTS .RTQGDWTGS IESEKKVESK
98BWMO14_1  GRKKRRRRR. .KAPPSSEDH QNLVSKQPLP .QSRGNPTGS EESKKKVESK
98BWMO18_d  GRKKRRQRR. .SAPPSSKDH QNPIPKQPLP .QTRGDQTGS EESKKKVESK
98BWMO36_a  GRKKRSQRR. .STPPSSEDH QDPISKQPLP QQPRGIPTGS EESKKKVESK
98BWMO37_d  GRKKRRQRR. .SAPPSSEDH QNLISKQPLS .QTRGDPTGP EESKKKVESK
99BW3932_1  GRKKRRPRR. .SAPPSNKDH QDLVSKQPLP .QTRGDPTGS EESKKKVESK
99BW4642_4  GRKKRGQRRR .STPSSSKDH QDLVPKQPLP .RTQGNSTGS EESKKKVESK
99BW4745_8  GRKKRRQRR. .SAPPSSEDH QNLVSKQPLP .QAQGDPTGS EESKKKVESK
99BW4754_7  GRKKRRQRR. .SAPPSSTDH QTPISKQPLP .QTRGDQTGS EKSKKKVESK
99BWMC16_8  GRKKRRQRR. .SAPPSSEDH QNPISKQPLP .RTQGDSTGS EESKKKVESK
A2_CD___97  GRKKRRPRR. .GPDQGNTDH QNPVPKQSLP .QTQRDSTGP EESKKKVESK
A2_CY___94  GRKKRRPRR. .KPSPSNKDH QNPIPKQSLP .QAQRVPTGP EEPKKEVESK
A2D___97_9  GRKKRRPRR. .RPPQGGTGH QNPIPEQSLP .RPQRIPTGT EESKTEVESK
A2G_CD___9  GRKKRKHRR. .GAPPGSKDH QNPIPKQPLP TTRG.NPTGP KESKKEVASK
A_BY_97_97  GRKKRRHRR. .GTSHSSKDH XIPISKQPLP HTQK.DQTGX EESKKKXESK
A_KE_93_Q2  GRKKRRQRR. .GAPQSSKDH QNSIQKQPIP QTQG.VSTGP KESKKKMESK
A_SE_93_SE  GRKKRTRRR. .GAPQSNKDH QNPISKQPIP QTQG.DSTGP EESKKKVESK
A_SE_94_SE  GRKKRRQRR. .RTPQSSKDH QNPIPKQPIP QTQG.VSTGP EESKKKVESK
A_SE_94_SE  GRKKRKQRR. .GAPQSNKDH QNPIQQQSIP QTQG.IPTGP EESKKKVESK
A_SE_95_SE  GRKKRKQRR. .GTPQGNKDH QNPIPKQSTP QTQG.ISTGP EESKKKVESK
A_SE_95_SE  GRKKRRQRR. .RTPQSSKDH QNPVPKQSIP QAQG.DSTGP EESKKKVESK
A_SE_95_UG  GRKKRRQRR. .GTPQSSKDH QDPIPKQSIP PAQG.IPTGP EESKKKVESK
A_UG_85_U4  GKKKRKPRR. .GPPQGSKDH QTLIPKQPLP QSQR.VSAGQ EESKKKVESK
A_UG_92_92  GRKKRKPRR. .GTPQSNKDH QNPIPKQPIP RTQG.DSTGP EESKKKVESK
AC_IN_95_2  GRKKWRQRR. .RAPPSSEDH QNLISKQPIP .QTQGDSTGP EESKKKMESK
AC_RW_92_9  GRKKRRQRR. .NAPPSSEDH QNPISKQPLS .QTRGDPTGS EESKKKVESK
AC_SE_96_S  GRKKRRQRR. .SAPPSSKDH QNPISKQSIP .QTQGDPTGP EEPKKKVESK
ACD_SE_95_  GRKKRRQRR. .GPPQSNKDH QNPVQKQPIP QTQR.ESTGP EESKKKVESK
ACG_BE___V  GRKKRSQRR. .RAPPSSEDR QNLISKQPLS .QTRGDPTGP EKSKKEVESK
AD_SE_93_S  GRKKRRQRR. .XTPEGGQAH QDPIPKQPSS QPRGD.PTGP KEKKK.VASK
AD_SE_95_S  GRKKRKQRR. .GAPPSSKDH QNPIPKQPIP QTQG.ISTGP KESKKEVESK
ADHK_NO_97  GRKKRRPRR. .RPPKSSQDH QDFIPKQPLS .RTHGEPTGP KEKKK.VASK
ADK_CD_85_  GRKKRRQRR. .RPPQGNQAH QDPLPEQPSS QHRGDHPTGP KEKKK.VESK
AG_BE___VI  GRKNRKHRR. .GTPQGSKDH QNPVPKQSLP LIRG.IPTGP EESKKEVASK
AG_NG_92_9  GRKKRRRRR. .GTPQSHQDH QNPVPKQPLP TTRG.NPTGP KESKKEVESK
AGHU_GA___  GRKKRSQRR. .RAPKSSPDH QNLVPKQPFS .RTNGNPTGP KEKKK.VASK
AGU_CD_76_  GRKKRRQRR. .GTPQDRKDH QNPVPRQPLP TTRG.NPTGP KESKKEVESK
AJ_BW_98_B  GRKKRRQRR. .TAPPGNKNH QDLVQEQPLS .QTQRKSTGP EESKKEVESK
B_AU___VH_  GRKKRRQRR. .RAPEDSQTH QVSLSKQSAP QPRGD.PTGP KESKKKVESK
B_CN___RL4  GRKKRRQRR. .RAPQDSQTH QASLSKQPAS QPRGD.PAGP KESKKKVESE
B_DE_86_D3  GRKKRRQRR. .RAPEDSQTH QVSLSKQPAS QPRGD.PTGP KESKKKVETE
B_DE_86_HA  GRKKRRQRR. .RAPQDSQTH QVSLPKQPSS QQRGD.PDSP KKSKKKVERE
B_FR_83_HX  GRKKRRQRR. .RAHQNSQTH QASLSKQPTS QPRGD.PTGP KE.KKKVERE
B_GA___OYI  GRKKRRQRR. .RAPQDSKTH QVSLSKQPAS QPRGD.PTGP KESKKKVERE
B_GB___CAM  GRKKRRQRR. .RTPQSSKTH QASLSKQPAS QFQGD.PTGP KESKKKVEGE
B_GB___GB8  GRKKRRQRR. .RLPEDSQIH QVSLPKQPTS QPQGD.PTGP KESKKKVESK
B_GB_59_MA  GRKKRRQRR. .RAPPDSQTR QVSLSKQPTS QPRGD.PTGP EESKKKVERE
```

B_KR___WK_	GRKKRRQRR.	.RAPQDNKNH	QVSLSKQPTS	RARGD.PTGQ	EESKEKVEKE
B_NL_86_32	GRKKRRQRR.	.RSPQDSETH	QVSLSKQPAS	QPRGD.PTGP	KESKKKVERE
B_TW___TWC	GRKKRRQRR.	.RTPQNSQTH	QADLSKQPTS	QPRGD.QTGQ	KESTKKVERE
B_US___DH1	GRKKRRKRR.	.RSPQHSQTD	QASLSKQPAS	QPRGD.PTGP	KESKKKVETE
B_US___P89	GRKKRRQRR.	.RPPQDSQTH	QVSLSKQPSS	QPRGD.PTGP	KEQKKKVERE
B_US_83_RF	GRKKRRQRR.	.GPPQGSQTH	QVSLSKQPTS	QPRGD.PTGP	KESKEKVERE
B_US_83_SF	GRKKRRQRR.	.RAPQDSQTH	QASLSKQPAS	QSRGD.PTGP	TESKKKVERE
B_US_84_MN	GRKKRRQRR.	.RAPEDSQTH	QVSLPKQPAP	QFRGD.PTGP	KESKKKVERE
B_US_86_JR	GRKKRRQRR.	.RPPQDSQTH	QVSLPKQPSS	QQRGD.PTGP	KESKKKVERE
B_US_86_YU	GRKKRRQRR.	.RPPQDSQTH	QSSLSKQPTS	QLRGD.PTGP	KESKKKVERE
B_US_87_BC	GRKKRRQRR.	.RAPQDSQTH	QASLSKQPTS	QPRGD.PTGP	TESKKKVERE
B_US_88_WR	GRKKRRQRR.	.RAPPEGLTH	QVPLSKQPSS	QFRGD.PTGP	KESKKKVERE
B_US_90_WE	GRKKRRQRR.	.RSPQNSQTH	QDSLSKQPTS	QPRGD.PTGP	KESKKKVVRE
BF1_BR_93_	GRKKRRQRH.	.RTPQSSQLH	QDPVPKQPAS	QAQGN.PTGP	KESKKKVERE
C_BR_92_92	GRKKRRQRR.	.SAPPSSEDH	QNPIPKQPLP	.QTRGDQTGS	KESKKKVESQ
C_BW_96_96	GRKKRRQRR.	.SAPPSSEDH	QNPVSKQPLP	.QTRGDPTGS	EESKKKVESK
C_BW_96_96	GRKKRRQRR.	.SAPPSSKDH	QNPVSKQPLP	.QTRGDPTGL	EESKKKVESK
C_BW_96_96	GRKKRRQRR.	.SAPPSSEDH	QDLVPKQPLS	.QARGNPTSS	KESKKKVESK
C_BW_96_96	GRKKRGQRR.	.SAPPRSEDH	QNLISKQPLP	.RTQGDSTGS	EESKKKVESK
C_ET_86_ET	GRKKRRQRR.	.RAPQSSKDH	QNLISKQPLS	.HTRGDPTGS	EESKKKVESK
C_IN_93_93	GRKKRRQRR.	.SAPPSSEDH	QNLISKQPLP	.RTQGDPTGS	EESKKKVESK
C_IN_93_93	GRKKRRQRR.	.RAPQSSEDH	QNLISKQPLP	.RTQGDPTGS	EESKKKVESK
C_IN_93_93	GRKKRRQRR.	.SAPPSSEDH	QNLISKQPLP	.RTQGDPTGS	EESKKKVESK
C_IN_94_94	GRKKRRQRR.	.SAPQSSEDH	QDLISKQPLP	.RTQGDPTGS	EESKKKVESK
C_IN_95_95	GRKKRRQRR.	.SAPQSSEDH	QNPISKQPLP	.RTPGDPTGS	EESKKKVEGK
CRF01_AE_C	GRKKRKHRR.	.GPPPGSKDH	QNPIPKQPLP	TTRG.NPTGP	KESKKKVESK
CRF01_AE_C	GRKKRKHRR.	.GPSQDSKDH	QNSIPKQPLP	TSRG.NPTGP	KESKKEVAKK
CRF01_AE_C	GRKKRKHRR.	.GTPQGSKGH	QDPISKQPLP	IIRG.NPTGP	KESKKKVESK
CRF01_AE_T	GRKKRKHRR.	.GTPQSSKDH	QNPIPKQPLP	IIRG.NPTGP	KESKKEVESK
CRF01_AE_T	GRKKRKHRR.	.GTPQSRKDH	QHPIPEQPLS	IIRR.NPTDP	KESKKEVASK
CRF01_AE_T	GRKKRKHRR.	.GTPQSSKDH	QSPIPEQPLP	IIRG.NPTDP	KESKKEVASK
CRF01_AE_T	GRKKRKHRR.	.GTPQSRKDH	QYPIPEQPLP	IIRG.NPTDP	KESKKEVASK
CRF01_AE_T	GRKKRKHRR.	.GTPQSSKDH	QTPIRKQPPS	IIRGGNPTDP	KESKKEVASK
CRF01_AE_T	GRKKRKHRR.	.RTPQSSKDH	QYPIPEQPSP	IIRG.NPTDP	KESKKKVESK
CRF02_AG_F	GRKKRRRRR.	.GTPQSRQDH	QNPVPKQPLP	TTRG.IPTDP	KESKKEVASK
CRF02_AG_F	GRKKRXRRR.	.GTPQSRQDH	QNPVPKQPLP	TTRG.DPTDP	KESKKEVASK
CRF02_AG_G	GRKKRRRRR.	.GTPQSRQDR	QNPVSKQPLP	TTRG.NPTGP	KESKREVESK
CRF02_AG_N	GRKKRRRRR.	.GTPQSHQDH	QNPVSKQSLP	QTRG.DPTGP	KESKKEVESK
CRF02_AG_S	GRKKRKRRR.	.GTPQSRQDH	QNPVPKQPLP	TTRG.NPTDP	KESKKEVESK
CRF02_AG_S	GRKKRRRRR.	.GTPQSRQDN	QDPVPKQPLP	TTRG.NPAGP	KESKKEVAGK
CRF03_AB_R	GRKKRRQRR.	.GTPQSRQDH	QNPVPKQPLP	TTRG.EQTGP	KESKKEVASK
CRF03_AB_R	GRKKRRQRR.	.RAPQDNQTD	QVSLPKQPAS	QPRGD.PTGP	KE.KKKMERE
CRF04_cpx_	GRKKRRQRR.	.RPPQDNQTD	QVSLPKQPAS	QPRGD.PTGP	KE.KKKVERE
CRF04_cpx_	GRKKRKHRR.	.GSLQGSKGH	QNLIPKQPLS	QQPNGDSTGP	EEQKKKVASK
CRF04_cpx_	GRKKRKRNE.	.DLLGFSRDR	QNPIPKQPLS	Q.PNGNPEGP	KEQKKKVASK
CRF04_cpx_	GRKKRKHRR.	.RPPQGSRDR	QNPIPKQPLS	QQHSGDPTGP	KEQKEAVASK
CRF05_DF_B	GRKKRRPRR.	.RPPQGSQAH	QDPVPEQPPS	QPRGD.PTGP	KKQKKEVESK
CRF05_DF_B	GRKKRRSRR.	.RPPQGGQAH	QIPVPEQPSS	QARGD.PTGQ	KEQKKKVESK
CRF06_cpx_	GRKKRRQRR.	.QAPPGSKNH	QDPVSKQPLS	.QTQREQTGP	EKSKKEVESK
CRF06_cpx_	GRKKRRQRR.	.TAPPGSKNH	QDPVPKQPLS	.QTQRGPTGP	EKSKKKVESK
CRF06_cpx_	GRKKRRQRR.	.TAPPGSKNH	QDPVPKQPLS	.QTQRKSTGP	EESKKEVESK
CRF11_cpx_	GRKKRRQRR.	.TAPLGSKSH	QDPVPKQPLS	.QTQRESTGP	EKSKEEVESK
CRF11_cpx_	GRKKRRQRR.	.AASHSSENH	QDPIPKQPST	.QPNRKPTGP	EESKKEVESK
CRF11_cpx_	GRKKWRQRR.	.TASRSSKNH	QDPIPEQPLP	.QASRNPTGP	EEPKKEVESK
D_CD_83_EL	GRKKRRQRR.	.GPPQGGQAH	QVPIPKQPSS	QPRGD.PTGP	KEQKKKVESE
D_CD_83_ND	GRKKRRQRR.	.KPPQGDQAH	QVPIPEQPSS	QSRGD.PTGP	K.KKKKVESE
D_CD_84_84	GRKKRRQRR.	.RPPHSSQTH	QDPIPKQPSS	QPRGD.PTGQ	KEKKK.VESK
D_UG_94_94	GRKKRRPRR.	.RTPPGGQAN	QDPVPKQPSS	QPRGN.PTGP	KEKKK.VESE
F1_BE_93_V	GRKKRRQRH.	.RTPQSSQVH	QNSLPKQPLS	QARGD.PTGP	KESKKEVESK
F1_BR_93_9	GRKKRRQRP.	.RTPQSSQIH	QDFVPKQPIS	QARGN.PTGP	KESKKEVESK
F1_FI_93_F	GRKKRRQRH.	.RTPQSSQIH	QDPVPKQPLS	QPRRN.PTGP	KESKKEVESK
F1_FR_96_M	GRKKRRQRR.	.RTPQSSQSH	KNPIPEQPLS	QARGD.PTGP	KESKKEVESK

```
F2_CM_95_M   GRKKRRQRR. .RTPQSGEVH QDPVSKQPLS QTRGD.PKGP EESKKKVESK
F2KU_BE_94   GRKKRRQRR. .RTPQSSQAH QNPISKQPLS QARGD.PTGP KEPKKEVESK
G_BE_96_DR   GRKKRKHRR. .GTPHSSKDH QTPVPKQPFS TTRG.NPTGP QESKKEVESK
G_NG_92_92   GRKKRRPRR. .GTPQGSKDH QNPVPKQPLP ITSG.NPTGS EKPKKEVASK
G_SE_93_SE   GRKKRKHRR. .GTPQSSKGH QDPVPKQPLP TTRG.NPTGP KESKKEVASK
H_BE___VI9   GRKKRRQRR. .GTPKSLQDH QTLIPKQPLS .RTSGDPTGP EKKKK.VASK
H_BE___VI9   GRKKRSRRR. .ATPASVQDH QNHIPKQPLS .RTRGDPTGP KEKKK.VASK
H_CF_90_90   GRKKRSQRH. .RTPASLQDH QNSISKQPLS .RTHGDPTGP KEQKKEVASK
J_SE_93_SE   GRKKRRQRR. .SAPPGSKTH QDLIPKQPLS .QTQRKPTGP EESKKEVESK
J_SE_94_SE   GRKKRRQRR. .SAPPGSKNH QDLIPEQPLF .QTQRKPTGP EESKKEVESK
K_CD_97_EQ   GREKRRQRT. .TTPYASKNH KDPIPKQPLP .QARGDPTGP KESKKEVESK
K_CM_96_MP   GRKKRRPRR. .TTPYNSENH QDPLRKQPLS .QPRGEQTDP KESKKKVESK
N_CM_95_YB   GRKKRSQRR. .RTPQSSKSH QDLIPEQPLS .QQQGDQTGQ KKQKEALESK
O_CM___ANT   GRKK...RGR PAAAS.HPDH KDPVPKQSPT ITK.RKQERQ EEQEEEVEKK
O_CM_91_MV   GRKK...RRR PAAAASYPDN KDPVPEQSLS HTG.RKQKRQ EEQEKKVEKE
O_SN___99S   GRKK...RRR PAAAARHPDN QDIVPEQLTY ITN.RKQKRQ EEQEKEVENE
O_SN___99S   GRKK...RRR PAAAARNPDN QDIVPEQPPP ITNNRKHKRQ EEQEKEVEKE
U_CD___83C   GRKKRGKRR. .RTPQSGPNH QNIVSKQPSS QPRGD.PTGQ EEPKKKVEKK

             101  108
00BW0762_1   TETDPFD.
00BW0768_2   TKTDQFD.
00BW0874_2   TKTDQFD.
00BW1471_2   TEADPCD.
00BW1616_2   TETDPFD.
00BW1686_8   TKTDPFD.
00BW1759_3   TETDRFD.
00BW1773_2   TETDPD..
00BW1783_5   TETDPFD.
00BW1795_6   TETDPFD.
00BW1811_3   TETDPD..
00BW1859_5   TETDPYD.
00BW1880_2   TETNPFD.
00BW1921_1   TEADQFD.
00BW2036_1   TEADRFD.
00BW2063_6   TETDPFD.
00BW2087_2   TERDPFD.
00BW2127_2   TTTDPFD.
00BW2276_7   TETDPYD.
00BW3819_3   TKTDPFD.
00BW3842_8   TETDRFD.
00BW3871_3   TKTDQFD.
00BW3876_9   TKADPFD.
00BW3886_8   AETDQFDY
00BW3891_6   TETDPFA.
00BW3970_2   TERDPFA.
00BW5031_1   TETDPFDW
96BW01B21    TKTDPFD.
96BW0407     TEADPFD.
96BW0502     TEADPFA.
96BW06_J4    TETDQFD.
96BW11_06    TETDQFD.
96BW1210     TETDPFD.
96BW15B03    TETDRFD.
96BW16_26    TETDPCD.
96BW17A09    TEADPFD.
96BWMO1_5    TKTDQFD.
96BWMO3_2    TETDPFD.
98BWMC12_2   TKAHPFD.
98BWMC13_4   TETDQFD.
98BWMC14_a   TDTDQFA.
```

98BWMO14_1	TETDPCA.
98BWMO18_d	TETDQFD.
98BWMO36_a	TETDPFD.
98BWMO37_d	TETDPFD.
99BW3932_1	TETDPFD.
99BW4642_4	TETDQFA.
99BW4745_8	TEPDPCD.
99BW4754_7	TETDPFD.
99BWMC16_8	TEADRFD.
A2_CD___97	AETDRFD.
A2_CY___94	AETDRFD.
A2D___97_9	AETDPCD.
A2G_CD___9	TETDPD..
A_BY_97_97	AETDQFD.
A_KE_93_Q2	AEADRFD.
A_SE_93_SE	AETDRFD.
A_SE_94_SE	AEADRFD.
A_SE_94_SE	AETDRFD.
A_SE_95_SE	TEADRFD.
A_SE_95_SE	TETDRFA.
A_SE_95_UG	AETDRFA.
A_UG_85_U4	AKTDRFA.
A_UG_92_92	TEADRYA.
AC_IN_95_2	AKTDRFD.
AC_RW_92_9	TEADPFD.
AC_SE_96_S	TETDRFD.
ACD_SE_95_	AETDRFD.
ACG_BE___V	TETHPLA.
AD_SE_93_S	AEADQFDW
AD_SE_95_S	TEPDRFD.
ADHK_NO_97	TXTDPFDW
ADK_CD_85_	AEADQFDW
AG_BE___VI	TETHPGD.
AG_NG_92_9	TETDQCA.
AGHU_GA___	AEADPFDW
AGU_CD_76_	TETDPFAW
AJ_BW_98_B	AKPDRFD.
B_AU___VH_	TETNPSD.
B_CN___RL4	TETDPRD.
B_DE_86_D3	TETDPID.
B_DE_86_HA	TEADPFD.
B_FR_83_HX	TETDPFD.
B_GA___OYI	TETDPED.
B_GB___CAM	TETHPGD.
B_GB___GB8	TETDPSDW
B_GB_59_MA	TETDPVA.
B_KR___WK_	TVVDPVT.
B_NL_86_32	TETDPVD.
B_TW___TWC	TETDPNDQ
B_US___DH1	TETDPVH.
B_US___P89	TETDPVH.
B_US_83_RF	TETDPAVQ
B_US_83_SF	TETDPFD.
B_US_84_MN	TETHPVD.
B_US_86_JR	TETDPDN.
B_US_86_YU	TETDPVH.
B_US_87_BC	TETDPVD.
B_US_88_WR	TETDPIA.
B_US_90_WE	TETDPED.
BF1_BR_93_	AKTDPD..
C_BR_92_92	TETDPFD.

```
C_BW_96_96  TETDPFD.
C_BW_96_96  TETDQFD.
C_BW_96_96  TETDPFD.
C_BW_96_96  TETDRFD.
C_ET_86_ET  AETDPYA.
C_IN_93_93  TKTDPFD.
C_IN_93_93  AKTDPFA.
C_IN_93_93  TKTDPFA.
C_IN_94_94  TTSDPFD.
C_IN_95_95  TKTDPFD.
CRF01_AE_C  AKTDPFA.
CRF01_AE_C  AETDPDW.
CRF01_AE_C  TKTDPCA.
CRF01_AE_T  AETDQCD.
CRF01_AE_T  AETDPCD.
CRF01_AE_T  AETDPCD.
CRF01_AE_T  AETDPCD.
CRF01_AE_T  AETDPD..
CRF01_AE_T  AETDQCD.
CRF02_AG_F  TETDQGD.
CRF02_AG_F  TKTDPCD.
CRF02_AG_G  TETDPFA.
CRF02_AG_N  TKTDPCD.
CRF02_AG_S  TETDPCD.
CRF02_AG_S  TETGPCD.
CRF03_AB_R  TETHPFD.
CRF03_AB_R  TETHPFD.
CRF04_cpx_  TEADPFA.
CRF04_cpx_  TEADPFD.
CRF04_cpx_  TESNPFD.
CRF05_DF_B  TEADQFDW
CRF05_DF_B  AETDPFDC
CRF06_cpx_  AEPDRFD.
CRF06_cpx_  AEPDRFD.
CRF06_cpx_  AETDRFD.
CRF06_cpx_  TEPDRFD.
CRF11_cpx_  AEPDRFD.
CRF11_cpx_  AEPAPFD.
D_CD_83_EL  AETDPDC.
D_CD_83_ND  AETDPFDW
D_CD_84_84  TEVHPFDW
D_UG_94_94  TEADPFDW
F1_BE_93_V  AKTDPCA.
F1_BR_93_9  AKTDPD..
F1_FI_93_F  AKTDPCD.
F1_FR_96_M  TETDPFD.
F2_CM_95_M  TKTDPSD.
F2KU_BE_94  TETDPLD.
G_BE_96_DR  TETDPFD.
G_NG_92_92  TETDPLD.
G_SE_93_SE  AEADQCD.
H_BE___VI9  TETDPFDW
H_BE___VI9  TEADPCD.
H_CF_90_90  TETDPD..
J_SE_93_SE  AEPDRFD.
J_SE_94_SE  AEPDRFD.
K_CD_97_EQ  TKTDPD..
K_CM_96_MP  TKTDQFD.
N_CM_95_YB  TEADPCD.
O_CM___ANT  AGPGGYPR
O_CM_91_MV  TGPSGQPC
```

```
O_SN___99S  ACP.RYPG
O_SN___99S  TGSDRYPR
U_CD___83C  TTTDPFD.
```

Table 17. HIV Vif Sequence Alignment
GCG Multiple Sequence File.
Written by Omiga 1.1

Name: 00BW0762_1	Len:	194	Check:	4675	Weight:	1.00
Name: 00BW0768_2	Len:	194	Check:	4961	Weight:	1.00
Name: 00BW0874_2	Len:	194	Check:	3755	Weight:	1.00
Name: 00BW1471_2	Len:	194	Check:	3843	Weight:	1.00
Name: 00BW1616_2	Len:	194	Check:	4613	Weight:	1.00
Name: 00BW1686_8	Len:	194	Check:	4096	Weight:	1.00
Name: 00BW1759_3	Len:	194	Check:	3523	Weight:	1.00
Name: 00BW1773_2	Len:	194	Check:	4446	Weight:	1.00
Name: 00BW1783_5	Len:	194	Check:	3151	Weight:	1.00
Name: 00BW1795_6	Len:	194	Check:	4892	Weight:	1.00
Name: 00BW1811_3	Len:	194	Check:	3877	Weight:	1.00
Name: 00BW1859_5	Len:	194	Check:	3290	Weight:	1.00
Name: 00BW1880_2	Len:	194	Check:	2555	Weight:	1.00
Name: 00BW1921_1	Len:	194	Check:	4284	Weight:	1.00
Name: 00BW2036_1	Len:	194	Check:	4019	Weight:	1.00
Name: 00BW2063_6	Len:	194	Check:	4165	Weight:	1.00
Name: 00BW2087_2	Len:	194	Check:	5068	Weight:	1.00
Name: 00BW2127_2	Len:	194	Check:	5231	Weight:	1.00
Name: 00BW2128_3	Len:	194	Check:	5469	Weight:	1.00
Name: 00BW2276_7	Len:	194	Check:	5547	Weight:	1.00
Name: 00BW3819_3	Len:	194	Check:	1251	Weight:	1.00
Name: 00BW3842_8	Len:	194	Check:	4197	Weight:	1.00
Name: 00BW3871_3	Len:	194	Check:	3487	Weight:	1.00
Name: 00BW3876_9	Len:	194	Check:	4432	Weight:	1.00
Name: 00BW3886_8	Len:	194	Check:	5175	Weight:	1.00
Name: 00BW3891_6	Len:	194	Check:	3845	Weight:	1.00
Name: 00BW3970_2	Len:	194	Check:	2268	Weight:	1.00
Name: 00BW5031_1	Len:	194	Check:	3711	Weight:	1.00
Name: 96BW01B21	Len:	194	Check:	4602	Weight:	1.00
Name: 96BW0407	Len:	194	Check:	5108	Weight:	1.00
Name: 96BW0502	Len:	194	Check:	4385	Weight:	1.00
Name: 96BW06_J4	Len:	194	Check:	5371	Weight:	1.00
Name: 96BW11_06	Len:	194	Check:	6037	Weight:	1.00
Name: 96BW1210	Len:	194	Check:	4343	Weight:	1.00
Name: 96BW15B03	Len:	194	Check:	5690	Weight:	1.00
Name: 96BW16_26	Len:	194	Check:	4471	Weight:	1.00
Name: 96BW17A09	Len:	194	Check:	3907	Weight:	1.00
Name: 96BWMO1_5	Len:	194	Check:	5608	Weight:	1.00
Name: 96BWMO3_2	Len:	194	Check:	3079	Weight:	1.00
Name: 98BWMC12_2	Len:	194	Check:	5336	Weight:	1.00
Name: 98BWMC13_4	Len:	194	Check:	5304	Weight:	1.00
Name: 98BWMC14_a	Len:	194	Check:	3984	Weight:	1.00
Name: 98BWMO14_1	Len:	194	Check:	2480	Weight:	1.00
Name: 98BWMO18_d	Len:	194	Check:	2801	Weight:	1.00
Name: 98BWMO36_a	Len:	194	Check:	3762	Weight:	1.00
Name: 98BWMO37_d	Len:	194	Check:	4971	Weight:	1.00
Name: 99BW3932_1	Len:	194	Check:	4165	Weight:	1.00
Name: 99BW4642_4	Len:	194	Check:	2912	Weight:	1.00
Name: 99BW4745_8	Len:	194	Check:	5323	Weight:	1.00
Name: 99BW4754_7	Len:	194	Check:	3964	Weight:	1.00
Name: 99BWMC16_8	Len:	194	Check:	6325	Weight:	1.00
Name: A2_CD_97CD	Len:	194	Check:	5849	Weight:	1.00
Name: A2_CY_94CY	Len:	194	Check:	5097	Weight:	1.00
Name: A2D___97KR	Len:	194	Check:	3871	Weight:	1.00
Name: A2G_CD_97C	Len:	194	Check:	5705	Weight:	1.00
Name: A_BY_97BL0	Len:	194	Check:	8467	Weight:	1.00
Name: A_KE_Q23_A	Len:	194	Check:	5053	Weight:	1.00

Name: A_SE_SE659	Len:	194	Check:	3808	Weight:	1.00
Name: A_SE_SE725	Len:	194	Check:	5856	Weight:	1.00
Name: A_SE_SE753	Len:	194	Check:	5873	Weight:	1.00
Name: A_SE_SE853	Len:	194	Check:	5523	Weight:	1.00
Name: A_SE_SE889	Len:	194	Check:	3207	Weight:	1.00
Name: A_SE_UGSE8	Len:	194	Check:	5837	Weight:	1.00
Name: A_UG_92UG0	Len:	194	Check:	5055	Weight:	1.00
Name: A_UG_U455_	Len:	194	Check:	5386	Weight:	1.00
Name: AC_IN_2130	Len:	194	Check:	3540	Weight:	1.00
Name: AC_RW_92RW	Len:	194	Check:	3664	Weight:	1.00
Name: AC_SE_SE94	Len:	194	Check:	4187	Weight:	1.00
Name: ACD_SE_SE8	Len:	194	Check:	4653	Weight:	1.00
Name: ACG_BE_VI1	Len:	194	Check:	6680	Weight:	1.00
Name: AD_SE_SE69	Len:	194	Check:	6416	Weight:	1.00
Name: AD_SE_SE71	Len:	194	Check:	8542	Weight:	1.00
Name: ADHK_NO_97	Len:	194	Check:	1255	Weight:	1.00
Name: ADK_CD_MAL	Len:	194	Check:	5519	Weight:	1.00
Name: AG_BE_VI11	Len:	194	Check:	7396	Weight:	1.00
Name: AG_NG_92NG	Len:	194	Check:	7120	Weight:	1.00
Name: AGHU_GA_VI	Len:	194	Check:	5827	Weight:	1.00
Name: AGU_CD_Z32	Len:	194	Check:	4744	Weight:	1.00
Name: AJ_BW_BW21	Len:	194	Check:	4938	Weight:	1.00
Name: B_AU_VH_AF	Len:	194	Check:	6911	Weight:	1.00
Name: B_CN_RL42_	Len:	194	Check:	6101	Weight:	1.00
Name: B_DE_D31_U	Len:	194	Check:	3568	Weight:	1.00
Name: B_DE_HAN_U	Len:	194	Check:	6199	Weight:	1.00
Name: B_FR_HXB2_	Len:	194	Check:	4714	Weight:	1.00
Name: B_GA_OYI__	Len:	194	Check:	4534	Weight:	1.00
Name: B_GB_CAM1_	Len:	194	Check:	4796	Weight:	1.00
Name: B_GB_GB8_A	Len:	194	Check:	6277	Weight:	1.00
Name: B_GB_MANC_	Len:	194	Check:	4800	Weight:	1.00
Name: B_KR_WK_AF	Len:	194	Check:	3856	Weight:	1.00
Name: B_NL_3202A	Len:	194	Check:	4181	Weight:	1.00
Name: B_TW_TWCYS	Len:	194	Check:	5670	Weight:	1.00
Name: B_US_BC_L0	Len:	194	Check:	4644	Weight:	1.00
Name: B_US_DH123	Len:	194	Check:	5023	Weight:	1.00
Name: B_US_JRCSF	Len:	194	Check:	6235	Weight:	1.00
Name: B_US_MNCG_	Len:	194	Check:	2067	Weight:	1.00
Name: B_US_P896_	Len:	194	Check:	6322	Weight:	1.00
Name: B_US_RF_M1	Len:	194	Check:	5045	Weight:	1.00
Name: B_US_SF2_K	Len:	194	Check:	3723	Weight:	1.00
Name: B_US_WEAU1	Len:	194	Check:	4222	Weight:	1.00
Name: B_US_WR27_	Len:	194	Check:	7503	Weight:	1.00
Name: B_US_YU2_M	Len:	194	Check:	5093	Weight:	1.00
Name: BF1_BR_93B	Len:	194	Check:	4341	Weight:	1.00
Name: C_BR_92BR0	Len:	194	Check:	5265	Weight:	1.00
Name: C_BW_96BW0	Len:	194	Check:	5846	Weight:	1.00
Name: C_BW_96BW1	Len:	194	Check:	3799	Weight:	1.00
Name: C_BW_96BW1	Len:	194	Check:	4343	Weight:	1.00
Name: C_BW_96BW1	Len:	194	Check:	5690	Weight:	1.00
Name: C_ET_ETH22	Len:	194	Check:	4205	Weight:	1.00
Name: C_IN_93IN1	Len:	194	Check:	3033	Weight:	1.00
Name: C_IN_93IN9	Len:	194	Check:	3201	Weight:	1.00
Name: C_IN_93IN9	Len:	194	Check:	4905	Weight:	1.00
Name: C_IN_94IN1	Len:	194	Check:	3129	Weight:	1.00
Name: C_IN_95IN2	Len:	194	Check:	3351	Weight:	1.00
Name: CRF01_AE_C	Len:	194	Check:	6355	Weight:	1.00
Name: CRF01_AE_C	Len:	194	Check:	2596	Weight:	1.00
Name: CRF01_AE_C	Len:	194	Check:	4412	Weight:	1.00
Name: CRF01_AE_T	Len:	194	Check:	5882	Weight:	1.00
Name: CRF01_AE_T	Len:	194	Check:	5558	Weight:	1.00

```
Name: CRF01_AE_T       Len:   194  Check: 5926  Weight:   1.00
Name: CRF01_AE_T       Len:   194  Check: 5579  Weight:   1.00
Name: CRF01_AE_T       Len:   194  Check: 2960  Weight:   1.00
Name: CRF01_AE_T       Len:   194  Check: 5867  Weight:   1.00
Name: CRF02_AG_F       Len:   194  Check: 1879  Weight:   1.00
Name: CRF02_AG_F       Len:   194  Check: 3893  Weight:   1.00
Name: CRF02_AG_G       Len:   194  Check: 5632  Weight:   1.00
Name: CRF02_AG_N       Len:   194  Check: 3187  Weight:   1.00
Name: CRF02_AG_S       Len:   194  Check: 5274  Weight:   1.00
Name: CRF02_AG_S       Len:   194  Check: 5177  Weight:   1.00
Name: CRF03_AB_R       Len:   194  Check: 5215  Weight:   1.00
Name: CRF03_AB_R       Len:   194  Check: 5211  Weight:   1.00
Name: CRF04_cpx_       Len:   194  Check: 2914  Weight:   1.00
Name: CRF04_cpx_       Len:   194  Check: 5450  Weight:   1.00
Name: CRF04_cpx_       Len:   194  Check: 4358  Weight:   1.00
Name: CRF05_DF_B       Len:   194  Check: 7168  Weight:   1.00
Name: CRF05_DF_B       Len:   194  Check: 5710  Weight:   1.00
Name: CRF06_cpx_       Len:   194  Check: 4977  Weight:   1.00
Name: CRF06_cpx_       Len:   194  Check: 5603  Weight:   1.00
Name: CRF06_cpx_       Len:   194  Check: 4458  Weight:   1.00
Name: CRF06_cpx_       Len:   194  Check: 3711  Weight:   1.00
Name: CRF11_cpx_       Len:   194  Check: 4246  Weight:   1.00
Name: CRF11_cpx_       Len:   194  Check: 7186  Weight:   1.00
Name: D_CD_84ZR0       Len:   194  Check: 4173  Weight:   1.00
Name: D_CD_ELI_K       Len:   194  Check: 5080  Weight:   1.00
Name: D_CD_NDK_M       Len:   194  Check: 4285  Weight:   1.00
Name: D_UG_94UG1       Len:   194  Check: 3203  Weight:   1.00
Name: F1_BE_VI85       Len:   194  Check: 5281  Weight:   1.00
Name: F1_BR_93BR       Len:   194  Check: 2780  Weight:   1.00
Name: F1_FI_FIN9       Len:   194  Check: 3522  Weight:   1.00
Name: F1_FR_MP41       Len:   194  Check: 3777  Weight:   1.00
Name: F2_CM_MP25       Len:   194  Check: 5402  Weight:   1.00
Name: F2KU_BE_VI       Len:   194  Check: 6170  Weight:   1.00
Name: G_BE_DRCBL       Len:   194  Check: 6155  Weight:   1.00
Name: G_NG_92NG0       Len:   194  Check: 5616  Weight:   1.00
Name: G_SE_SE616       Len:   194  Check: 6641  Weight:   1.00
Name: H_BE_VI991       Len:   194  Check: 5850  Weight:   1.00
Name: H_BE_VI997       Len:   194  Check: 6598  Weight:   1.00
Name: H_CF_90CF0       Len:   194  Check: 4443  Weight:   1.00
Name: J_SE_SE702       Len:   194  Check: 6028  Weight:   1.00
Name: J_SE_SE788       Len:   194  Check: 5724  Weight:   1.00
Name: K_CD_EQTB1       Len:   194  Check: 6926  Weight:   1.00
Name: K_CM_MP535       Len:   194  Check: 6479  Weight:   1.00
Name: N_CM_YBF30       Len:   194  Check: 4619  Weight:   1.00
Name: O_CM_ANT70       Len:   194  Check: 412   Weight:   1.00
Name: O_CM_MVP51       Len:   194  Check: 6622  Weight:   1.00
Name: O_SN_99SE_       Len:   194  Check: 8844  Weight:   1.00
Name: O_SN_99SE_       Len:   194  Check: 9492  Weight:   1.00
Name: U_CD___83C       Len:   194  Check: 5631  Weight:   1.00

            1                                                  50
00BW0762_1  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSKRANGWFY RHHYESRHPK
00BW0768_2  MENRWQGLIV WQVDRMKIRT WNSLVKHHMY VSRRANGWFY RHHYESRHPK
00BW0874_2  MENRWQVLIV WQVDRMKIRA WNSLVKHHMY ISRKASGWFY RHHYESRHPK
00BW1471_2  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISRRAKGWVY RHHYESRHPR
00BW1616_2  MENRWQVLIV WQVDRMRIRT WNSLVKHHMY VSRRASGWFY RHHYESRHPK
00BW1686_8  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISRRASGWSY RHHYESRHPK
00BW1759_3  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISKRAKGWLY RHHYENRHPK
00BW1773_2  MENRWQVLIV WQVDRMKIKT WNSLVKHHMY VSKRAKGWFY RHHYESSHPR
00BW1783_5  MENRWQVLIV WQVDRMRIRT WNSLVKHHMY ISKKARGWFY RHHYESRHPK
00BW1795_6  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSRKANGWFY RHHYESRHPK
```

```
00BW1811_3  MENRWQVLIV WQVDRMKIKT WNSLVKHHMY ISKKAKGWFY RHHYESRNPK
00BW1859_5  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISRKAKGWYY RHHFESRHPK
00BW1880_2  MENRWQVLIV WQIDRMKIRT WNSLVKHHMY ISRRASGWFY RHHYESRNPK
00BW1921_1  MENRWQVLIV WQIDRMKIRT WNSLVKHHMY ISRRANGWFY RHHYESRHPK
00BW2036_1  MENRWQALIV WQVDRMRIRT WNSLVKHHMH VSKRAKGWFY RHHFESRHPK
00BW2063_6  MENRWQGLIV WQVDRMRIRT WNSLVKHHMY ISRRASGWFY RHHYDSRHPK
00BW2087_2  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISRRAAGWFY RHHYESRNPR
00BW2127_2  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY TSKKATGWFY RHHYESRHPK
00BW2128_3  MENRWQVLIV WQVDRMRIRT WNSIVKHHMY VSRRTNGWFY KHHYESRNPK
00BW2276_7  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISRRTMGWFY RHHFQSRHPK
00BW3819_3  MENRWQVLIV WQVDRMKIRT WNSLVKHHMH ISKRAKGWFY RHHFESRHPK
00BW3842_8  MENRWQALIV WQVDRMRIRT WNSLVKHHMY ISRRASGWFY RHHFESRHPK
00BW3871_3  MENRWQVLIV WQVDRMRIRT WNSLVKHHMY ISRRASGWFY RHHYESRHPK
00BW3876_9  MENRWQVLIV WQVDRMRIRT WNSLVKHHMY VSRKAHGWFY RHHYQSRHPK
00BW3886_8  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISKRANGWFY RHHYQSRHPK
00BW3891_6  MENRWQVMIV WQVDRMKIRT WNSLVKHHMY VSKKANGWFY RHHYESRHPR
00BW3970_2  MENRWQVLIV WQIDRMRIKT WNSLVKHHMY VSRRASGWFY RHHFESRHPK
00BW5031_1  MDNRWQGLIV WQVDRMRIRT WNSLVKHHMY VSRRANGWFY RHHHESRHPK
 96BW01B21  MENRWQVLIV WQVDRMRIRT WNSLVKHHMY VSRRASGWFY RHHFESRHPK
  96BW0407  MENRWQVMIV WQVDRMKIRT WNSLVKHHMY VSKKAKGWFY RHHYESRHPR
  96BW0502  MENRWQVLIV WQVDRMKIRT WNSLVKHHMH ISKRAKGWFY RHHYESRHPK
 96BW06_J4  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISKRANGWFY RHHYESRHPK
 96BW11_06  MENRWQALIV WQVDRMRIRT WTSLVKHHMY VSRRANGWYY RHHYESRHPK
  96BW1210  MENRWQGLIV WQVDRMRIRT WHSLIKHHMY VSKRADGWFY RHHYESRHPK
 96BW15B03  MENRWQALIV WQVDRMRIRT WNSLVKHHMY VSKRTNGWFY RHHFESRHPK
 96BW16_26  MENRWQVLIV WQVDRMKIKT WNSLVKHHMY ISRRANGWSY GHHYESRNPK
 96BW17A09  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISRRAKGWFY RHHYESRHPK
 96BWMO1_5  MENRWQGLIV WQVDRMKIRT WNSLVKHHMY VSKRAAGWWY RHHYESRHPK
 96BWMO3_2  MENRWQVLIV WQVDRMRIRT WNSLVKHHMY VSKKAAGWFY RHHYESRHPK
98BWMC12_2  MENRWQVLIV WQVDRMRIRT WNSLVKHHMY TSGRASGWFY RHHYESRHPK
98BWMC13_4  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSKRAKGWYY RHHYESRHPK
98BWMC14_a  MENRWQGLIV WQVDRMKIRT WNSLVKHHMY ISRRASGWFY RHHFESRHPK
98BWMO14_1  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY ISRRAKGWIY KHHFESRNPK
98BWMO18_d  MENRWQVLIV WQVDRMRIRT WNSLVKHHMY VSKRAKGWFY RHHYESRHPK
98BWMO36_a  MENRWQVLIV WQVDRMRIRA WNSLVKHHMH ISKRAAGWFY RHHYESRNPK
98BWMO37_d  MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSKRASKWFY RHHYESRHPK
99BW3932_1  MENRWQVLIV WQVDRMRIRT WNSLVKHHMY ISRRAEGWFY RHHYESRHPK
99BW4642_4  MENRWQVLIV WQIDRMKIRT WNSLVKHHMY VSKRAKGWFY RHHFESRHPK
99BW4745_8  MENRWQVLIV WQVDRMRIRT WNSLVKHHMH ISRRANGWFY RHHYESRHPR
99BW4754_7  MENRWQVMIV WQVDRMKIRT WNSLVKHHMY VSRRANGWFY RHHYESRHPK
99BWMC16_8  MENRWQGLIV WQVDRMRIRT WNSLVKHHMH VSRRANGWFY RHHYESRHPK
A2_CD_97CD  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY VSKKAREWFY RHHYESRHPR
A2_CY_94CY  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISRKAKGWVY KHHYESRNPR
A2D___97KR  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY VSKKAKDWCY RHHYESINPR
A2G_CD_97C  MENRWQVMIV WQVDRMRIKR WNSLVKHHMY VSRKAKDWFY RHHYESRHPK
A_BY_97BL0  XENRWQVXIV WQVDRMRIRT WNSLVKHHMY VSKKAR.EVY RHHYESRQPR
A_KE_Q23_A  MENRWQAMIV WQVDRMRIRT WNSLVKHHMH VSKKAKRWFY RHHYESRHPK
A_SE_SE659  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY VSKKAKNWVY RHHFESRHPK
A_SE_SE725  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY VSRKAKDWFY RHHYESRNPR
A_SE_SE753  MENRWQVMIV WQVDRMRIRT WNSLVKHHMC VSKKARNWFY RHHYESRHPK
A_SE_SE853  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISKKAKNWFY RHHFESRHPK
A_SE_SE889  MENRWQVMVV WQVDRMRIRT WNSLVKHHMY ISKKAKGWLY RHHFESRHPK
A_SE_UGSE8  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISKKAAGWFY RHHYESRHPK
A_UG_92UG0  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISRRAKGWFY RHHYESRHPK
A_UG_U455_  MENRWQVMIV WQVDRMKIRT WNSLVKHHMY VSKKAQGWFY RHHYESRHSR
AC_IN_2130  MENRWQALIV WQVDRMKIRT WNSLVKHHMY VSRKANGWFY RHHYDSRHPK
AC_RW_92RW  MENRWQVMIV WQVDRMKIRT WNSLVKHHMY ASRRAKGWFY RHHYESRHPK
AC_SE_SE94  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISKKAKRWFY RHHYESRHPK
ACD_SE_SE8  MENRWQVMIV WQVDRMRIGT N.SLVKHHMY VSKKARGWFY RHHYXTRHPR
ACG_BE_VI1  MENRWQVMVV WQVDRMRIRT WHSLVKHHMY TSKKAKNWCY RHHYESMHPK
AD_SE_SE69  MENRWQVMIV WQVDRMRIRT WKSLVKYHMY VSKQARGWLY RHHYDCLNPK
```

```
AD_SE_SE71   MENRWQVMIV WQVDRMRIKT WNSLVKHHMY VSKKAQNWVY RHHYESRHPR
ADHK_NO_97   MENRWQVMIV WQVDRMRIRT WHSLVKHHIY VSKKANKWLF RHHYESRHPK
ADK_CD_MAL   MENRWQVMIV WQVDRMRIRT WHSLVKHHMY VSKKAKNWFY RHHYESRHPK
AG_BE_VI11   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY VSKKAKGWFY RHHYESRHPK
AG_NG_92NG   MENRWQVVIV WQVDRMRIRT WNSLVKYHMY KSKKAKDWFY RHHYESRHPK
AGHU_GA_VI   MENRWQVMIV WQVDRMRIST WKSLVKHHMY VSKKAQGWFY RHHYDCTHPR
AGU_CD_Z32   MENRWQVMIV WQVDRMRINT WKGLVKYHMY KSKKAKNWFY RHHYDSNHPK
AJ_BW_BW21   MENRWQVMIV WQVDRMRINT WKSLVKYHMH VSKKTKKWLY RHHYDSNHPK
B_AU_VH_AF   MENRWQVMIV WQVDRMRIRT WKSLVKHHLY KSGKARRWVY RHHYESTHPR
B_CN_RL42_   MENRWQVMIV WQVDRMRIKT WKSLVKHHMY ISRKAKGWFY KHHYDSTHPK
B_DE_D31_U   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSGKAEKWFY KHHYESTNPR
B_DE_HAN_U   MENRWAVMIV WQVDRMRIRT WNSLVKHHIY CSRKAKNWVY RHHYESTNPR
B_FR_HXB2_   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSGKARGWFY RHHYESPHPR
B_GA_OYI__   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSKKAKGWFY RHHYESTHPR
B_GB_CAM1_   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY ISGKAKKWSY RHHYESTHPR
B_GB_GB8_A   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY ISGKAKKWVY KHHYENTHPR
B_GB_MANC_   MESRWQVMIV WQVDRMRIRT WKSLVKHHMY ISGKAKRWSY KHHYESTNPR
B_KR_WK_AF   MENRWQVMIV WQVDRMRIKT WKSLVKHHMY ISKKAKEWVY RHHYESTHPR
B_NL_3202A   MENRWQVMIV WQVDRMRIRA WKSLVKHHMY KSKKAERWFY RHHYESTHPR
B_TW_TWCYS   MENRWQVMIV WQVDRMRIRA WKSLVKHHMY ISKKAKGWLY KHHYESTHPR
B_US_BC_L0   MENRWQVMIV WQVDRMRIRT WISLVKHHMY ISRKAKGWFY RHHYESTHPK
B_US_DH123   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY VSKKAKGWFY RHHYESTHPR
B_US_JRCSF   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISGKAKGWIY KHHYESTNPR
B_US_MNCG_   MENRRQVMIV WQADRMRIRT WKSLVKHHMY ISKKAKGRFY RHHYESTHPR
B_US_P896_   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY ISGKAKGWSY RHHYESTNPR
B_US_RF_M1   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY ISRKAKGWFY RHHYESTHPR
B_US_SF2_K   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY ISKKAKGWFY RHHYESTHPR
B_US_WEAU1   MENRWQVMIV WQVDRMRIRT WKSLVKHHMY ISKKAKGWSY RHHYESTHPR
B_US_WR27_   MENRWQVMIV WQVDRMRIRT WKSLVKHHXH ISGKARRWXY XHHYENNHPR
B_US_YU2_M   MENRWQVMIV WQVDRMRIRA WKSLVKHHMY ISGKARGWFY RHHYESPHPR
BF1_BR_93B   MENRWQVVIV WQVDRMRINT WKSLVKYHMH VSKKAKRWFY RHHFESRHPR
C_BR_92BR0   MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSRRASGWYY RHHYESRHPK
C_BW_96BW0   MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSRKANGWFY RHHYESRHPR
C_BW_96BW1   MENRWQVLIV WQVDRMRIRT WTSLVKHHMY VSRRANGWSY RHHFESRHPK
C_BW_96BW1   MENRWQGLIV WQVDRMRIRT WHSLIKHHMY VSKRADGWFY RHHYESRHPK
C_BW_96BW1   MENRWQALIV WQVDRMRIRT WNSLVKHHMY VSKRTNGWFY RHHFESRHPK
C_ET_ETH22   MENRWQVLIV WQVDRMKIRT WNSLVKHHMH ISRRANGWVY RHHYDSRHPK
C_IN_93IN1   MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSRRAKGWFY RHHYDSRHPK
C_IN_93IN9   MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSRRANGWFY RHHYESRHPK
C_IN_93IN9   MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSRRATGWFY RHHYESRNPK
C_IN_94IN1   MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSRRANGWFY RHHYDSRNPK
C_IN_95IN2   MENRWQVLIV WQVDRMKIRT WNSLVKHHMY VSRRANGWFY RHHYESRHPK
CRF01_AE_C   MENRWQVMIV WQVDRMRIRT WYSLVKHHMY ISKKAKNWFY RHHYESQHPK
CRF01_AE_C   MENRWQVMIV WQVDRMRIRA WNSLVKHHMY SSKKAAKWFY RHHYESQHPK
CRF01_AE_C   MENRWQVMIV WQVDRMRIKT WNSLVKHHMY ISKKAKKWVY RHHYESQHPK
CRF01_AE_T   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISKKAKKWFY RHHYESQHPK
CRF01_AE_T   MENRWQVMIV WQVDRMRIRT WNSIVKHHMY ISKKAKKWFY RHHYESQHPK
CRF01_AE_T   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISKKAKQWFY RHHYESQHPK
CRF01_AE_T   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY VSKKAKKWFY RHHYESQNPK
CRF01_AE_T   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISKKAKKWFY RHHYESQHPK
CRF01_AE_T   MENRWQVMIV WRVDRMRIRT WNSLVKHHMY ISKKAKNWFY RHHYESQHPK
CRF02_AG_F   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISRKAKNWFY RHHFESRHPK
CRF02_AG_F   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISRKAKNWFY RHHFESRHPK
CRF02_AG_G   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY VSRKAKDWFY RHHYESRHPK
CRF02_AG_N   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY VSKKAKGWFY RHHFESRHPK
CRF02_AG_S   MENRWQVMIV WQVDRMRIRT WNSLVKYHMY KSRKAKDWFY RHHYESSHPR
CRF02_AG_S   MENRWQVMIV WQVDRMRIRT WNSLVKHHMY VSKKAHRWFY RHHYESRHPK
CRF03_AB_R   MENRWQVMIV WQVDRMRIRT WNSLVKHHIY ISKKARGWVY KHHYESRNPR
CRF03_AB_R   MENRWQVMIV WQVDRMRIRT WNSLVKHHIY ISKKARGWVY KHHYESRNPR
CRF04_cpx_   MANRWQVMIV WQVDRMKIRT WNSLVKHHMY VSKKAKGF.Y RHHYESRHPK
CRF04_cpx_   MENRWQVMIV WQVDRMKIRT WNSLVKHHMY ISKKAKGWSY RHHYESRHPR
```

```
CRF04_cpx_  MENRWQVMTV WQVDRMKIRT WNSLVKHHMH ISKKAKGWVY KHHYESRNPR
CRF05_DF_B  MENRWQVMIV WQVDRMRINT WKSLVKYHMH VSKKANRWCY RHHFESRNPR
CRF05_DF_B  MENRWQVMIV WQVDRMRINT WKSLVKYHMH VSKKTKAWFY RHHYESRHPR
CRF06_cpx_  MENRWQVMIV WQVDRMRINT WKSLVKYHMH ISKKAKRWNY RHHYDSNHPK
CRF06_cpx_  MENRWQVMIV WQVDRMRINT WKSLVKYHMH ISKKAKRWTY RHHYDSNHPK
CRF06_cpx_  MENRWQVMIV WQVDRMRINA WKSLVKYHMN VSKKAKGWLY RHHYDSNHPK
CRF06_cpx_  MENRWQVMIV WQVDRMRINT WKSLVKYHMH ISKKARKWAY RHHYDSHHPK
CRF11_cpx_  MENRWQVMIV WQVDRMRIRT WHSLVKHHMY VSKKARRWMY RHHYESRHPK
CRF11_cpx_  MENRWQVMIV WQVARMRIRT WNSLVKHHMY VSKKAKGWLY RHHYESRHPR
D_CD_84ZR0  MENRWQVMIV WQVDRMRINT WKSLVKYHMH ISKKAKGWFY RHHYDSPHPK
D_CD_ELI_K  MENRWQVMIV WQVDRMRIKT WKSLVKHHMY VSKKANRWFY RHHYESPHPK
D_CD_NDK_M  MENRWQVMIV WQVDRMRINT WKSLVKYHMY VSKKANRWFY RHHYDSHHPK
D_UG_94UG1  MENRWQVMIV WQVDRMRIRT WKSLVKHHMY ISKKAKGWLY RHHYDCPNPK
F1_BE_VI85  MENRWQLMIV WQVDRMRINT WKSLVKYHMY VSKKAKGWSY RHHFQSRHPR
F1_BR_93BR  MENRWQVMIV WQVDRMRINT WKSLVKYHMH ISKKAKGWFY RHHFESRHPK
F1_FI_FIN9  MENRWQVTIV WQVDRMRINT WKSLVKYHMH VSKKAKRWFY RHHFESRHPK
F1_FR_MP41  MENRWQVMIV WQVDRMRIST WKSLVKYHMH VSKKAKNWFY RHHFQSRHPK
F2_CM_MP25  MENRWQVMIV WQVDRMKIRT WNSLVKHHMY ISKRAAGWFY RHHYESRHPR
F2KU_BE_VI  MENRWQVMIV WQVDRMRIKT WNSLVKHHMY VSRKTKGWSY RHHYESINPR
G_BE_DRCBL  MENRWQVMIV WQVDRMRIRT WNSLVKHHMY ISKKAKGWFY RHHYESRHPK
G_NG_92NG0  MENRWQVVIV WQVDRMRIRT WNSLVKHHMY VSKKAKGWFY RHHYESRHPR
G_SE_SE616  MENRWQVMIV WQVDRMRIRT WHSLVKHHMY VSKKARGWFY RPHYASRHPR
H_BE_VI991  MENRWQVMIV WQVDRMRIKT WNSLVKHHMY VSKKAKKWVY RHHYESTNPK
H_BE_VI997  MENRWQVMIV WQVDRMRINK WKSLVKYHMY ISKKANRWVY RHHYDSPHPR
H_CF_90CF0  MENRWQVMIV WQVDRMRINT WKSLVKYHMH ISRKARGWFY RHHFESTHPR
J_SE_SE702  MENRWQVMIV WQVDRMRINT WKSLVKYHMN VSKKARQWLY RHHYDSRHPK
J_SE_SE788  MENRWQVMIV WQVDRMRINT WKSLVKYHMN VSKKARKWLY RHHYDSNHPK
K_CD_EQTB1  MENRWQVMIV WQVDRMRINT WKSLVKYHMY VSKKANRWRY RHHYDSNHPK
K_CM_MP535  MENRWQVMIV WQVDRMRINT WKSLVKYHMH ISKKANRWYY RHHYESRHPK
N_CM_YBF30  MENRWQVMVV WQVDRMKIRK WNSLVKHHMY VSKKAKGWYY RHHYETHHPK
O_CM_ANT70  MENRWQVLIV WQVDRQKVKA WNSLVKYHKY RSRKTENWWY RHHYESRNPR
O_CM_MVP51  MENRWQVLIV WQIDRQKVKA WNSLVKYHKY MSKKAANWRY RHHYESRNPK
O_SN_99SE_  MENRWQVLIV WQVDRQRVKT WNSLVKYHKY RSGKTRDWYY RHHFESRNPR
O_SN_99SE_  MENRWQVLIV WQVDRQRVKT WNSLVKYHKY RSGKTRDWYY RHHFESKNPR
U_CD___83C  MENRWQVMIV WQVDRMRIKT WNSLVKHHMY ISKKAKGWVY KHHYESTNPR

            51                                                  100
00BW0762_1  LSSEVHIPLG D.ARLVIKTY WGLHTGERDW HLGHGVSIEW KLRRYSTQVD
00BW0768_2  VSSEVHIPLG D.ARLVIITY WGLHTGERDW HLGHGVSIEW RLRRYSTQVD
00BW0874_2  ISSEVHIPLG D.ARLVIKTY WGLQTGEREW HLGHGVSIEW RLRRYSTQVD
00BW1471_2  ISSEVHIPLG D.AKLVIKTY WGLHTGEREW HLGHGVSIEW RLKSFSTQVD
00BW1616_2  ISSEVHIPLG D.ARLIIKTY WGLQTGEKDW HLGHGVSIEW RLRKYSTQVD
00BW1686_8  ISSEVHIPLG E.ARLVITTY WGLQTGEREW HLGHGVSIEW RLRRYSTQVD
00BW1759_3  ISSEVHIPLG D.ARLVITTY WGLNTGERDW HLGHGVSIEW RLGRYSTQVD
00BW1773_2  VSSEIHIPLG E.ARLVIKTY WGLQTGEREW HLGHGVSIEW RWKRYSTQVE
00BW1783_5  ISSEVHIPLG D.ARLVIKTY WGLHTGEREW HLGHGVSIEW RLRKYSTQID
00BW1795_6  ISSEVHIPLG E.ARLVIITY WGLQTGEREW HLGHGVSIEW RLRRYSTQVD
00BW1811_3  ISSEVHIPLG D.ARLVIKTY WGLQTGERDW HLGHGVSIEW RLRKYSTQVE
00BW1859_5  ISSEVHIPLG E.ARLVVKTY WGLQTGEREW HLGHGVSIEW RLRGYSTQVD
00BW1880_2  ISSEVHIPLG E.ARLVIKTY WGLHTGERDW HLGHGVSIEW KLKRYSTQVD
00BW1921_1  VSSEVHIPLG E.ARLVITTY WGLQTGEREW HLGHGVSIEW RLRRYKTQVD
00BW2036_1  VSSEVHIPLG E.ARLVVITY WGLQTGEREW HLGQGVSIEW RLRRYSTQVD
00BW2063_6  ISSEVHIPLG D.ARLVIKTY WGLQTGERDW HLGHGVSIEW RLRKYSTQVD
00BW2087_2  VSSEVHIPLG D.ARLVIKTY WGLQTGERAW HLGHGVSIEW RMKRYSTQVD
00BW2127_2  VSSEVHIPLG E.ARLVIKTY WGLQTGERDW HLGHGASIEW RLRRYSTQVV
00BW2128_3  VSSGVHIPLG D.ARLVIKTY WGLQTGEREW HLGHGVSIEW RLRKYSTQVE
00BW2276_7  VSSEVHIPLG D.ARLVITTY WGLQTGEREW HLGHGVSIEW RLKRYSTQVD
00BW3819_3  VSSEVHIPLG D.ARLVIKTY WGLQTGERDW HLGHGVSIEW RLGKYNTQVE
00BW3842_8  VSSEVHIPLG E.ARLVIITY WGLQTGEREW HLGHGVSIEW RLRRYSTQVD
00BW3871_3  VGSEVHIPLG E.ARLVIRTY WGLQTGEREW HLGHGVSIEW RLRKYSTQVD
00BW3876_9  VSSEVHIPLG E.DKLVITTY WGLQTGEREW HLGHGVSIEW RLKKYSTQVD
```

```
00BW3886_8   VSSEVHIPLG E.ARLVITTY WGLQTGEREW HLGHGVSIEW RLKRYSTQVD
00BW3891_6   ISSEVHIPLG E.ARLIIKTY WGLQTGERDW HLGHGASIEW RLRRYSTQVD
00BW3970_2   ISSEVHIPLG E.ARLIIKTY G.LQTGEGDW HLGHGCSIEW RLRKYSTQVD
00BW5031_1   VSSEVHIPLG D.ARLVIKTY WGLHTGERDW HLGHGVSIEW RLRRYNTQVD
 96BW01B21   VSSEVHIPLG E.ARLVIKTY WGLQTGEREW HLGHGVSIEW RLRRYSTQVD
  96BW0407   VSSEVHIPLG D.ARLVIKTY WGLQTGERDW HLGHGVSIEW RLRGYSTQVD
  96BW0502   VSSEVHIPLG E.ARLVIKTY WGLQTGERDW HLGHGVSIEW RLRKYSTQVE
 96BW06_J4   VSSEVHIPLG D.ARLVITTY WGLQTGEREW HLGHGVSIEW RLRRYSTQVD
 96BW11_06   VSSEVHISLG E.ARLVIKTY WGLHTGERDW HLGHGVSIEW KLRGYSTQVD
  96BW1210   VSSEVHIPLG D.ARLVIITY WGLQTGEREW HLGHGVSIEW RLRSYSTQVD
 96BW15B03   VSSEVHIPLG E.ARLVIITY WGLQT.EREW HLGHGVSIEW RLRRYSTQVD
 96BW16_26   VSSEVHIPLG D.AKLVIKTY WGLQTGERDW HLGHGVSIEW RWGNYSTQVE
 96BW17A09   ISSEVHIPLG D.AKLVITTY WGLHTGEGDW HLGHGVSIEW RLKRFSTQVD
 96BWMO1_5   ISSEVHIPLG D.ARLVIKTY WGLNTGERDW HLGHGVSIEW RLRKYSTQVD
 96BWMO3_2   VSSEIHIPLG D.ARLIVRTY WGLQTGEKEW QLGHGVSIEW RLRSFSTQVD
98BWMC12_2   VSSEVHIPLG K.ARLVITTY WGLQIGERDW HLGHGVSIEW RLRKYSTQVD
98BWMC13_4   ISSEVHIPLG D.ARLVIKTY WGLNTGERDW HLGHGVSIEW RLRKYSTQVD
98BWMC14_a   VSSEVHIPLG E.AKLVIITY WGLQPGEREW HLGHGASIEW RLRRYSTQVD
98BWMO14_1   ISSEIHIPLG E.ARLVIKTY WGLNTGERDW HLGHGVSIEW RMRGYSTQVD
98BWMO18_d   ISSEVHIPLG D.AKLVIKTY WGLQTGEREW HLGHGTSIEW ILGGYSTQVD
98BWMO36_a   ISSEVHIPLG D.ARLVVKTY WGLQTGERDW HLGHGVSIEW RLRRYSTQVD
98BWMO37_d   VSSEVHIPLG D.ARLIIKTY WGLQTGERDW HLGHGCSIEW RLRRYSTQVD
99BW3932_1   VSSEVHIPLG E.ARLVIKTY WGLQTGEKDW HLGHGVSIEW RLKRYSTQVD
99BW4642_4   VSSEVHIPLG E.ARLIVRAY WGLQTGEREW HLGHGVSIEW RLRRYSTQVD
99BW4745_8   VSSEVHIPLG E.ARLVITTY WGLLPGEREW HLGHGVSIEW RLRRYSTQVD
99BW4754_7   ISSEVHIPLR D.ARLVIKTY WGLHTGERDW HLGHGVSIEW RLGRYSTQVD
99BWMC16_8   VSSEVHIPLG D.ARLVITTY WGLHTGEREW HLGHGVSIVW RLRRYSTQVD
A2_CD_97CD   VSSEVHIPLK E.ARLIVRTY WGLHPGEKDW HLGHGVSIEW RQGRYSTQID
A2_CY_94CY   ISSEVHIPLG E.ARIIVRTY WGLHIGEKDW HLGHGVSIEW RQNRYHTQID
A2D___97KR   VSSEVHIPLG G.AKIIVKTY WGLHPGEKDW HLGQGVSIEW RQERYSTQID
A2G_CD_97C   VSSEVHIPLG E.ATLVIRTY WGLQTGEKDW QLGHGVSIEW RQRKYSTQID
A_BY_97BL0   VSSEVHIPLG D.ARLVVKTY XGLHAXEKDW QLGHXVSIEX RQEXYSTQID
A_KE_Q23_A   VSSEVHIPLG D.ATLVVRAY WGLHTGEKDW HLGHGVSIEW RLKRYSTQIT
A_SE_SE659   VSSEVHIPLG D.AKLVVRTY WGLHTGEKEW HLGHGVSIEW RLNRYSTQID
A_SE_SE725   TSSEVHIPLG D.ARLVVRTY WGLHTGEKDW QLGHGVSIEW RLRRYSTQID
A_SE_SE753   VSSEVHIPLG D.ARLVVRTY WGLQTGEKDW HLGHGVSIEW RLKRYSTQID
A_SE_SE853   VSSEVHIPLG E.AKLVVRTY WGLQTGEKDW QLGHGVSIEW RLRRYSTQID
A_SE_SE889   VSSEVHIPLG E.ARLVVRTY WGLQTGEKDW HLGHGVSIEW RLRRYSTQID
A_SE_UGSE8   VSSEVHIPLG D.ARLIVRTY WGLHPGERDW QLGHGVSIEW RLRRYSTQID
A_UG_92UG0   VSSEVHIPIG D.ARIVVRTY WGLQTGEKDW HLGHGVSIEW RLKRYSTQID
A_UG_U455_   VSSEVHIPLG E.ARLVVRTY WGLHTGEKDW HLGHGVSIEW RLKRYSTQVD
AC_IN_2130   VSSEVHIPLG E.AKLVIKTY WGLQTGERDW HLGHGVSIEW RLRRYSTQVE
AC_RW_92RW   ISSEVHIPLG E.ARLVIKTY WGLQTGERDW HLGHGVSIEW RLRRYKTQVD
AC_SE_SE94   TSSEVHIPLG E.ARLVIITY WGLQTGERDW HLGHGVSIEW RLGRYRTQVD
ACD_SE_SE8   ISSEVHIPLG D.AKIVVRTY WGLHTGEKDW HLGHGVSIEW RLRKYSTQID
ACG_BE_VI1   VSSEVHIPLG D.ARIVVRTY WGLHTGEKAW QLGHGVSIEW RQRRYSTQID
AD_SE_SE69   VSSEVHIPLG E.ARLVVKTY WGLHTGERDW HLGQGVSIEW RKRRYSTQVD
AD_SE_SE71   VSSEVHIPLG D.AKLVVRTY WGLHTGEKDW HLGHGVSIEW RLRRYSTQID
ADHK_NO_97   VSSEVHIPLG D.ARLVVRTY WGLQTGEKAW HLGHGVSIEW RQKRYSTQID
ADK_CD_MAL   VSSEVHIPLG D.ARLVVRTY WGLQTGEKDW HLGHGVSIEW RQKRYSTQLD
AG_BE_VI11   VSSEVHIPLG D.ARLVVRTY WGLHTGEKDW HLGHGVSIEW RQRRYSTQID
AG_NG_92NG   VSSEVHIPLG E.ARLVVRTY WGLHTGERDW HLGQGVSIEW KQRRYSTQID
AGHU_GA_VI   VSSEVHIPLG D.ARLVIKTY WGLHTGERDW HLGQGVSIEW RKGRYSTQVD
AGU_CD_Z32   ISSEVHIPLE E.AKLVVTTY WGLHTGEREW HLGQGVSIEW RLRRYRTQVD
AJ_BW_BW21   ISSEVHIPLG S.AELVVTTY WGLNTGEREW HLGQGLSIEW RLKRYRTQVD
B_AU_VH_AF   ISSEVHIPLG E.ARLVITTY WGLHTGERDW HLGQGVSIEW RKRRYSTQVD
B_CN_RL42_   ISSEVHIPLG D.ARLVITTY WGLHTGERDW HLGQGVSIEW RKKRYSTQVD
B_DE_D31_U   ISSEVHIPLG D.ARLVITTY WGLHTGERDW QLGHGVSIEW RKKRYSTQVD
B_DE_HAN_U   VSSEVHIPLG E.AKLVITTY WGLHTGERDW HLGQGASIEW RKRRYSTQVD
B_FR_HXB2_   ISSEVHIPLG D.ARLVITTY WGLHTGERDW HLGQGVSIEW RKKRYSTQVD
B_GA_OYI__   ISSEVHIPLG D.ATLVVTTY WGLHTGEREW HLGQGASIEW RKKRYSTQVD
```

```
B_GB_CAM1_  ISSEVHIPLG E.ARLVVTTY WGLHTGERDW HLGQGVSIEW RTKGYNTQVD
B_GB_GB8_A  ISSEVHIPLG E.ARLVITTY WGLHTGERDW HLGQGVSIEW RKRRYRTQVD
B_GB_MANC_  VSSEVHIPLG D.AKLVIKTY WGLHTGERDW HLGQGASIEW RKKGYSIQVD
B_KR_WK_AF  ISSEVHIPLG D.AKLVITTY WGLHTGEREW HLGQGVSIEW RKKRYNTQVD
B_NL_3202A  ISSEVHIPVG E.ARLVITTY WGLHTGERDW HLGQGVSIEW RKKRYSTQVD
B_TW_TWCYS  ISSEVHIPLG D.ATLVITTY WGLHTGERDW HLGQGVSIEW RKRRYSTQVD
B_US_BC_L0  ISSEVHIPLG D.ARLVITTY WGLHTGERDW HLGHGVSVEW RKKRYSTQVD
B_US_DH123  ISSEVHIPLG D.ASLVVTTY WGLHTGERDW HLGQGVSIEW RKRRYSTQVD
B_US_JRCSF  VSSEVQIPLG D.ARLVITTY WGLHTGERDW HLGQGVSMEW RTRRYSTQVD
B_US_MNCG_  ISSEVHIPLG D.ARLVITTY WGLHTGERDW HLGQGVSIEW RKKRYSTQVD
B_US_P896_  ISSEVHIPLG D.AKLVVTTY WGLHTGERDW HLGQGVSIEW RKKRYSTQVD
B_US_RF_M1  ISSEVHIPPG D.ERLVITTY WGLHTGERDW HLGQGVSIEW RKRRYSTQVD
B_US_SF2_K  VSSEVHIPLG D.AKLVITTY WGLHTGEREW HLGQGVAIEW RKKKYSTQVD
B_US_WEAU1  ISSEVHIPLG E.GKLVITTY WGLHTGERDW HLGQGVSIEW RKQRYSTQVD
B_US_WR27_  ISSEVHIPLG D.AMLVITTY WGLHTGERDW HLGQGVSIEW RKQNYRTQVT
B_US_YU2_M  ISSEVHIPLG D.AKLVITTY WGLHTGERDW HLGQGVSIEW RKKRYSTQVD
BF1_BR_93B  VSSEVHIPLE E.AKLVITTY WGLHTGERDW HLGQGVSIEW RQGRYRTQID
C_BR_92BR0  ISSEVHIPLG E.ARLVIITY WGLQTGERDW HLGHGVSIEW RLRRYSTRVD
C_BW_96BW0  VSSEVHIPLG D.ARLVIKTY WGLQTGERDW HLGHGVSIEW RLRRYSTQVD
C_BW_96BW1  VSSEVHIPLG E.ARLVIKTY WGLHTGERDW HLGHGVSIEW TVRGYST.VD
C_BW_96BW1  VSSEVHIPLG D.ARLVIITY WGLQTGEREW HLGHGVSIEW RLRSYSTQVD
C_BW_96BW1  VSSEVHIPLG E.ARLVIITY WGLQT.EREW HLGHGVSIEW RLRRYSTQVD
C_ET_ETH22  VSSEVHIPLG E.ARLIIKTY WGLQTGERDW HLGHGVSIEW RLRSYNTQVD
C_IN_93IN1  VSSEVHIPLG E.ARLVIKTY WGLQTGERDW HLGHGVSIEW RLRRYNTQIE
C_IN_93IN9  VSSEVHIPLG E.ATLVIKTY WGLQTGERDW HLGHEVSIEW RLRRYNTQIE
C_IN_93IN9  ISSEVHIPLG E.ARLVIKTY WGLQTGERDW HLGHGVSIEW RLRRYSTQVE
C_IN_94IN1  VSSEVHIPLG E.AILVIKTY WGLQTGERDW HLGHGVSIEW RLRRYNTQIE
C_IN_95IN2  VSSEVHIPLG E.ARLVITTY WGLQTGERDW HLGHGVSIEW RLRKYSTQVE
CRF01_AE_C  VSSEVHIPIG D.ARLVIRTY WGLHTGEKDW HLGHGVSIEW RQRKYSTQVD
CRF01_AE_C  VSSEVHIPLG D.ARLIIRTY WGLHTGEKDW HLGHGVSIEW RQR.KSTQID
CRF01_AE_C  VSSEVHIPLG E.ARLVIRTY WGLHTGEKDW HLGHGVSIEW RQEKYSTQID
CRF01_AE_T  VSSEVHIPLG E.ARLVIRTY WGLQTGEKDW QLGHGVSIEW RQRNYSTQID
CRF01_AE_T  VSSEVHIPLG E.ARLVIRTY WGLQTGEKDW QLGHEVSIEW RQRTYSTQID
CRF01_AE_T  VSSEVHIPLG E.ARLVIRTY WGLQTGEKDW QLGHGVSIEW RQRKYSTQID
CRF01_AE_T  VSSEVHIPLE E.AKLIIRTY WGLQTGEKDW QLGHGVSIEW RQRTYSTQID
CRF01_AE_T  VSSEVHIPLG E.ARLVIRTY WGLQTGEKDW HLGHGVSIEW RQRKYSTQID
CRF01_AE_T  VSSEVHIPLG E.AKLVIRTY WGLQTGEKDW QLGHGVSIEW RQRKYSTQID
CRF02_AG_F  VSSEVHIPLG D.ARLIVRTY WGLHAGERDW HLGHGVSIEW KQRKYSTQID
CRF02_AG_F  VSSEVHIPLG D.ARLIVRTY WRLHAGERDW YLGHGVSIEW KQRKYSTQID
CRF02_AG_G  VSSEVHIPLG D.ARIVVRTY WGLHTGERDW HLGHGVSIEW RQKRYSTQID
CRF02_AG_N  VCSEVHIPLG D.ARLVVRTY WGLHTGERDW HLGHGVSIEW KQKRYSTQID
CRF02_AG_S  VSSEVHIPLG D.ARLVVRTY WGLHTGERDW HLGHGVSIEW KQRRYSTQID
CRF02_AG_S  VSSEVHIPLG D.ATLIVRTY WGLQPGERDW HLGHGVSIVW QQKRYSTQID
CRF03_AB_R  ISSEVHIPLG D.AKLVIKTY WGLHTGERDW HLGQGASIEW RKERYSTQVD
CRF03_AB_R  ISSEVHIPLG D.AKLVIKTY WGLHTGERDW HLGQGASIEW RKERYSTQVD
CRF04_cpx_  VSSEVHIPLG E.ARLVVRTY WGLQPGEQDW HLGHGVSIEW RLRRYSTQVD
CRF04_cpx_  VSSEVQIPLG D.ARLVIRTY WGLQPGEKDW HLGHGVSMEW RLRRYSTQVD
CRF04_cpx_  ASSEVHIPLG E.AKLVVRTY WGLQPGKKDW HLGHGVSIXW RLRSYSTQVD
CRF05_DF_B  ISSEVHIPLG D.AKLVVTTY WGLHTGERDW HLGQGVSIEW RKRRYSTQVD
CRF05_DF_B  ISSEVHIPLE E.AKLVIITY WGLHTGEREW HLGQGVSIEW RKGRYSTQID
CRF06_cpx_  ISSEVHIPLG S.AELVITTY WGLNTGERKW HLGQGVSIEW RLRRYRTQVD
CRF06_cpx_  VSSEVHIPLG C.AELVITTY WGLNTGERKW HLGQGVSIEW RLRRYRTQVD
CRF06_cpx_  ISSEVHIPLG C.AELVVTTY WGLNTGEREW HLGQGVSIEW RLKKYRTQVD
CRF06_cpx_  ISSEVHIPLG K.AELVVTTY WGLNTGERKW HLGQGVSIEW RLKRYRTQVD
CRF11_cpx_  ISSEVHIPLG D.ADLVVTTY WGLHTGEREW HLGQGVSIEW RMKRYRTQVD
CRF11_cpx_  VSSEVHIPLE E.DMLVVTTY WGLHTGEREW HLGQGVSIEW RRKRYRTQVD
D_CD_84ZR0  ISSEVHIPLG D.ARLVVTTY WGLHTGEREW HLGQGVSIEW RKRRYSTQVD
D_CD_ELI_K  ISSEVHIPLG E.ARLVIKTY WGLHTGEREW HLGQGVSIEW RKRRYSTQVD
D_CD_NDK_M  ISSEVHIPLG E.ARLVVTTY WGLHTGEKEW HLGQGVSIEW RKRRYSTQVD
D_UG_94UG1  ISSEVHIPLG E.ARLVVKTY WGLHTGEREW HLGQGVSIEW RKGRYNTQID
F1_BE_VI85  VSSEVHIPLE E.VKLVITTY WGLHPGEREW HLGQGVSIEW RQGKYRTQID
```

```
F1_BR_93BR   ISSEVHIPLE T.AELVITTY WGLLPGEREW HLGQGVSIEW RQGRYRTQID
F1_FI_FIN9   ISSEVHIPLE E.AELVITTY WGLNTGERDW HLGQGVSIEW RQGRYRTQID
F1_FR_MP41   ISSEVHIPLE E.AKLVVTTY WGLHTGERDW HLGQGVSIEW RQGRYRTQID
F2_CM_MP25   VSSEVHIPLE EDSKLVIITY WGLHTGERDW HLGQGVSIEW RQGRYRTQID
F2KU_BE_VI   VSSEVHIPLK E.AKLVVTTY WGLNTGERDW HLGQGVSIEW RQKRYRTQVD
G_BE_DRCBL   VSSEVHIPLG D.AKLVVRTY WGLHTGDKEW HLGHGVSIEW RLKKYSTQID
G_NG_92NG0   VSSEVHIPLR D.ATLVVRTY WGLHAGEKDW QLGHGVSIEW RQGRYSTQVD
G_SE_SE616   VSSEVHIPLG D.ATLVVTTY WGLHTGEKDW QLGHGVSIEW RQKRYSTQID
H_BE_VI991   TSSEVHIPVG D.ARLVITTY WGLHTGERDW HLGHGVSIEW RQRRYRTQVE
H_BE_VI997   ISSEVHIPLG D.ARLVITTY WGLHTGERDW HLGHGVSIEW RQERYSTQID
H_CF_90CF0   ISSEVHIPLG E.ARLVITTY WGLNTGEREW HLGQGVSIEW RQKRYSTQVD
J_SE_SE702   ISSEVHIPLG E.ARLVVTTY WGLQTGERDW HLGQGVSIEW RLKRYSTQVE
J_SE_SE788   ISSEVHIPLG E.AILVITTY WGLQTGERDW HLGQGVSIEW RRKRYRTQVD
K_CD_EQTB1   ISSEVHIPLG D.AELVVTTY WGLHTGERDW HLGQGVSIEW RQRRYRTQVD
K_CM_MP535   ISSEVHIPLG D.AELVVTTY WGLHTGEREW HLGQGVSIEW RLKKYRTQVD
N_CM_YBF30   ISSEVHIPVG Q.ARLVTVTY WGLLTGERDW HLGQGVSIEW RLKRYRTQVE
O_CM_ANT70   VSSSVYIPVG V.AHVVVTTY WGLMPGERDE HLGHGVSIEW RLRKYKTQVD
O_CM_MVP51   VSSAVYIPVA E.ADIVVTTY WGLMPGEREE HLGHGVSIEW RYKKYKTQID
O_SN_99SE_   VSSGVYIPVG G.PWIVVTTY WGLMPGERDE HLGHGVSIEW QYKEYKTQID
O_SN_99SE_   VSSGVYIPVG G.PWIVVTTY WGLMPGERDE HLGHGVSIEW RYKKYKTQID
U_CD___83C   ISSEVHIPLG N.ARIVVTAY WGLHTGERDW HLGQGVSIDW RYRKYTTQID
             ISSEVHIPLG N.ARIVVTAY WGLHTGERDW HLGQGVSIEW RQGRYSTQID

             101                                                  150
00BW0762_1   PGLADQLIHM HYFDCFADSA IRKALLGQVV IPRCDYQAGH NKVGSLQYLA
00BW0768_2   PGLADQLIHM HYFDCFADSA IRQAILGHIV IPRCDYQAGH TKVGSLQYLA
00BW0874_2   PGLADQLIHI HYFDCFADSA IRKAILGHIV TPRCDYQAGH NKVGSLQYLA
00BW1471_2   PGLADQLIHM HYFDCFAGSA IRKALLGQIV SPRCDYQAGH SKVGSLQYLA
00BW1616_2   PGLADQLIHM HYFDCFAESA IRKALLGQVV NPRCDYQAGH NKVGSLQYLA
00BW1686_8   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
00BW1759_3   PGLADQLIHM HYFDCFTDSA IRKALLGHIV IHRCDYQAGH NKVGSLQYLA
00BW1773_2   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYPAGH NKVGSLQYLA
00BW1783_5   PGLADQLIHM HYFDCFADSA IRKALLGHIV IPRCEYPAGH KKVGSLQYLA
00BW1795_6   PGLADQLIHT HYFDCFADSA IRKAILGHRV SPRCDYQAGH NKVGSLQYLA
00BW1811_3   PDLADQLIHI HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
00BW1859_5   PGLADQLIHM HYFDCFADSA IRKAILGHRV IHRCDYQAGH NKVGSLQYLA
00BW1880_2   PGLADL.IHM HYFDCFTDSA IRKAILGQVV IPRCDYQAGH NKVGSLQYLA
00BW1921_1   PGLADQLIHM HYFDCFAESA IRQAILGHIV IPRCDYQAGH NKVGSLQYLA
00BW2036_1   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
00BW2063_6   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
00BW2087_2   PGLADQLIHM HYFDCFADSA IRKAILGHIV SPRCDYQAGH NKVGSLQYLA
00BW2127_2   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCVYQAGH NKVGSIQYLA
00BW2128_3   PGLADQLIHM HYFDCFADSA IRRAILGHIV IPRCDYQAGH NKVGSLQYLA
00BW2276_7   PGLADQLIHM HYFDCFADSA IRQAILGHIV FPRCDYPAGH NKVGSLQYLA
00BW3819_3   PGLADQLIHM HYFDCFADSA IRKAILGHIV FPRCDYQAGH NKVGSLQYLA
00BW3842_8   PGLADQLIHI HYFDCFADSA IRKAILGHIV ISRCDYQAGH NKVGSLQYLA
00BW3871_3   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
00BW3876_9   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLT
00BW3886_8   PGLADQLIHM HYFDCFADSA LRKAILGHIV FPRCDYQAGH NKVGSLQYLA
00BW3891_6   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
00BW3970_2   PGLADQLIHM HYFDCFTESA IRKAILGHIV IPRCDYQAGH NQVGSLQYLA
00BW5031_1   PGLADQLIHM HYFDCFADSA IRKAILGRIV IPRCDYQAGH NQVGSLQYLA
 96BW01B21   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
  96BW0407   PGLADQLIHM HYFDCFADSA IRKAILGQIV SPRCEYQAGH NKVGSLQYLA
  96BW0502   PGLADQLIHM HYFDCFADSA IRKAILGEIV IPRCDYQAGH NQVGSLQYLA
 96BW06_J4   PGLADQLIHM YYFDCFADSA IRKAILGHIV NPRCDYQAGH NKVGSLQYLA
 96BW11_06   PGLADQLIHM HYFDCFTDSA IRKALLGQVV IPRCDYQAGH NKVGSLQYLA
  96BW1210   PGLADQLIHM HYFDCFAGSA IRQAILGHIV IPRCDYQAGH KKVGSLQYLA
 96BW15B03   PGLADQLIHM YYFDCFAESA IRKAILGHIV IPRCNYQAGH NKVGSLQYLA
 96BW16_26   PGLADQLIHM HYFDCFTDSA IRKAILGNKI IHRCNYPAGH NKVGSLQYLA
 96BW17A09   PGLADQLIHT HYFDCFADSA IRKAILGEIV SPRCDYPARH SQVGSLQYLA
 96BWMO1_5   PGLADQLIHM YYFDCFADSA IRRAILGYIV IPRCDYQAGH NKVGSLQYLA
```

```
 96BWMO3_2   PGLADQLIHM HYFDCFAESA IRKAILGHIV FPRCDYQAGH NKVGSLQYLA
98BWMC12_2   PGLADQLIHM HYFDCFTDSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
98BWMC13_4   PGLADQLIHM YYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
98BWMC14_a   PGLADRLIHM HYFDCFAESA IRNAILGHIV IPRCDYPAGH NQVGSLQYLA
98BWMO14_1   PGLADQLIHM HYFDCFADSA IRQAILGNLV IPRCDYQAGH NKVGSLQYLA
98BWMO18_d   PGLADQLIHM HYFDCFAESA IRKAILGDRV SPRCDYQAGH NKVGSLQYLA
98BWMO36_a   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYPAGH SKVGSLQYLA
98BWMO37_d   PGLADQLIHM HYFDCFADSA IRQAILGHIV IPRCDYQAGH NKVGSLQYLA
99BW3932_1   PSMADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGFLQYLA
99BW4642_4   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
99BW4745_8   PGMADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLKYLA
99BW4754_7   PGLADQLIHM HYFDCFTDSA IRKAILGHIV IHRCEYPAGH NKVGSLQYLA
99BWMC16_8   PSLADQLIHM HYFDCCADSA IRQAILGHIV IPRCNYPAGH NKVGSLQYLA
A2_CD_97CD   PDLADHLIHL YYFDCFSESA IRRAILGEIV RPRCEYQAGH NKVGSLQYLA
A2_CY_94CY   PDLADHLIHL YYFDCFSESA IRKAIIGEIV SPRCEYQAGH NKVGSLQYLA
A2D___97KR   PDLAGHLIHL HYFDCFSDSA IRKAILGKIV RPRCEYQAGH NKVGSLQYLA
A2G_CD_97C   PELADQLIHL HYFDCFSESA IRKAILGQVV RPRCQYQAGH TKVGSLQYLA
A_BY_97BL0   PDLADQLIHL YYFDCFSESA IRKAIVGHIV SPRCNYPAGH NKVGSLQYLA
A_KE_Q23_A   PDLADQLIHM HYFDCFSDSA IRKAIVGQVV SPKCEYQAGH NKVGSLQYLA
A_SE_SE659   PDQADQLIHL HYFDCFSDSA IRKALLGQVV SPRCEYQAGH NKVGSLQYLA
A_SE_SE725   PDLADQQIHL HYFDCFSDSA IRKAILGHVV SPICEYHTGH NKVGSLQYLA
A_SE_SE753   PDLADQLIHL YYLDCFSDSA IRKALLGQVV SPSCEYHTGH NQVGSLQYLA
A_SE_SE853   PDLADQLIHM HYXNCFSDSA IRKAILGQVV SPSCEYQAGH NKVGSLQYLA
A_SE_SE889   PDLADQLIHL HYFKCFSDSA IRKAILGEIV SPRCEYQAGH NKVGSLQYLA
A_SE_UGSE8   PDLADQLIHL HYFNCFSDSA IRKAILGRVV SPSCEYQTGH NKVGSLQYLA
A_UG_92UG0   PDLADQLIHL HYFNCFSDSA IRKAILGQVV SPRCDYQTGH NKVGSLQYLA
A_UG_U455_   PDLADHLIHL HYFDCFSESA IRRAILGQIV RPRCEYQAGH NKVGSLQYLA
AC_IN_2130   PGLADQLIHM HYFDCFADSA IRQAILGHIV IPRCDYQAGH .KVGSLQYLA
AC_RW_92RW   PGLAGQLIHM HYFDCFADSA IRKAILGHIV SPRCDYQAGH NKVGSLQYLA
AC_SE_SE94   PGLADQLIHM HYFDCFADSA IRKAILGHIV SPRCDYQAGH NKVGSLQYLA
ACD_SE_SE8   PDPADQLIHL HYFDCFSDSA IRKAILGQVV SPRCDYTAGH NKVGSLQYLA
ACG_BE_VI1   PDLADQLIHL YYFDCFSDSA IRKAILGQVV RPRCEYQAGH NKVGSLQYLA
AD_SE_SE69   PGLADQLIHI YYFDCFTESA IRKAILGHIV TTRCNYQTGH NKVGSLQYLA
AD_SE_SE71   PDQADQLIHL HYSNCFSESA IRKAILGQVV RPKCEYQTGH NKVGSLQYLA
ADHK_NO_97   PDLADHLIHL HYFDCFSDAV IRKAILGQVV RPRCEYQAGH NQVGSLQYLA
ADK_CD_MAL   PDLADQLIHL YYFDCFSESA IRQAILGHIV SPRCDYQAGH NKVGSLQYLA
AG_BE_VI11   PDLADQLIHL NYFDCFSDSA IRKAILGQVV RPRCEYQAGH NKVGSLQYLA
AG_NG_92NG   PDLADQLIHL HYFNCFSESA VRKAILGEVV RPRCEYQTGH NQVGSLQYLA
AGHU_GA_VI   PGLADQLIHM HYFDCFSDSA IRKAILGQVV RPRCEYSAGH NQVGSLQYLA
AGU_CD_Z32   PGLADQLIHM HYFDCFSESA IRKAILGHRV SPRCEYQAGH NKVGSLQYLA
AJ_BW_BW21   PGLADQLIHM HYFNCFSESA IRKAILGHIV SPICEYQAGH NKVGSLQYLA
B_AU_VH_AF   PGLADQLIHM YYFDCFSESA IRNAILERIV SPSCEHQAGH NKVGSLQYLA
B_CN_RL42_   PGLADQLIHL YYFDCFSESA IRNAILGRVV SPSCDYQAGH NKVGSLQYLA
B_DE_D31_U   PGLADQLIHL YYFDCFSESA IRNAILGRIG SPSCEYRAGH NKVGSLQYLA
B_DE_HAN_U   PNLADQLIHL YYFDCFSESA IRNAILGRIV SPRCEYQAGH SKVGSLQYLA
B_FR_HXB2_   PELADQLIHL YYFDCFSDSA IRKALLGHIV SPRCEYQAGH NKVGSLQYLA
B_GA_OYI__   PGLADQLIHT YYFDCFSESA IRNAILGNIV SPRCEYPAGH NKVGSLQYLA
B_GB_CAM1_   PDLADQLIHL YYFDCFSESA IRKAIVGRLV SPRCEYQAGH NKVGSLQYLA
B_GB_GB8_A   PGLADQLIHQ YYFDCFSESA IRNALLGRTV SPSCKYQAGH NKVGSLQYLA
B_GB_MANC_   PGLADQLIHL YYFDCFSESA IRNAILGHIV SPRCEYQAGH NKVGSLQYLA
B_KR_WK_AF   PDLADKLIHL HYFDCFSDSA IRHAILGHRV RPKCEYQAGH NKVGSLQYLA
B_NL_3202A   PGLADQLIHL YYFDCFSESA IRNAILGHVV SPRCEYQAGH NKVGSLQYLA
B_TW_TWCYS   PDQADQLIHL YYFDCFSESA IRKAIVGCRV SPRCEYQAGH NKVGSLQYLA
B_US_BC_L0   PDLADQLIHL YYFDCFSESA IRNAILGHIV SPRCEYQAGH NKVGSLQYLA
B_US_DH123   PDLADQLIHL YYFDCFSESA IRNAILGHRV SPRCEYQAGH NKVGSLQYLA
B_US_JRCSF   PDLADQLIHL YYFDCFSESA IRNAILGHIV SPRCEYQAGH SKVGSLQYLA
B_US_MNCG_   PDLADHLIHL HYFDCFSDSA IRKAILGHRV SPICEFQAGH NKVGPLQYLA
B_US_P896_   PGLADRLIHL YYFDCFSDSA IRKSILGHIV SPSCEYQAGH NKVGSLQYLA
B_US_RF_M1   PDLADQLIHL YYFDCFSESA IRKPSLGHIV SPRCEYQAGH NKVGSLQYLA
B_US_SF2_K   PGLADQLIHL HYFDCFSESA IKNAILGYRV SPRCEYQAGH NKVGSLQYLA
B_US_WEAU1   PDLADQLIHL YHFDCFSESA IRNAILGHLV IPRCEYQAGH NKVGSLQYLA
```

```
B_US_WR27_   PDLADQLIHR YYFDCFSEPA IRNTIVGRIV SPRCEYQTGH NKVGSLQYLA
B_US_YU2_M   PDLADQLIHL YYFDCFSESA IRKAILGYRV SPRCEYQAGH NKVGSLQYLA
BF1_BR_93B   PGLADQLIHI YYFDCFSESA IRKAILGHRI SPRCDYQAGH NKVGSLQYLA
C_BR_92BR0   PGLADQLIHM HYFDCFADSA IRKAILGHRV SSRCDYQAGH NKVGSLQYLA
C_BW_96BW0   PGLADQLIHM HYFDCFADSA IRKAILGQIV SPRCEYQAGH NKVGSLQYLA
C_BW_96BW1   PGLADQLIHM HYFDCFTDSA IRKALLGQVV IPRCDYQAGH NKVGSLQYLA
C_BW_96BW1   PGLADQLIHM HYFDCFAGSA IRQAILGHIV IPRCDYQAGH KKVGSLQYLA
C_BW_96BW1   PGLADQLIHM YYFDCFAESA IRKAILGHIV IPRCNYQAGH NKVGSLQYLA
C_ET_ETH22   PGLADHLIHM HYFDCFAESA IRKAILGYRV SPRCDYQAGH NKVGSLQYLA
C_IN_93IN1   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
C_IN_93IN9   PGLADQLIHM HYFDCFTDSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
C_IN_93IN9   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
C_IN_94IN1   PGLADQLIHM HYFDCFADSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
C_IN_95IN2   PGLADQLIHM HYFDCFADSA IRKAILGRIV IPRCDYQAGH NKIGSLQYLA
CRF01_AE_C   PDLADRQIHL QYFDCFSDSA IRKAILGHIV IPRCDYQAGH NKVGSLQYLA
CRF01_AE_C   PDLADQLIHL QYFDCFSDSA IRKAMLGQVV RPRCEYPTGH NKVGSLQYLA
CRF01_AE_C   PDLADQLIHL QYFDCFSDSA IRKALLGQIV RPRCEYPAGH NKVGSLQYLA
CRF01_AE_C   PDLADRLIHL QYFDCFSESA IRKAILGQVV RPRCDYPEGH NKVGSLQYLA
CRF01_AE_T   PDLADRLIHL QYFDCFSDSA IRRAILGQVV RRRCEYPSGH NKVGSLQYLA
CRF01_AE_T   PDLADQLIHL QYFDCFSDSA IRKAILGQVV RRRCEYPSGH NKVGSIQYLA
CRF01_AE_T   PDLADQLIHL QYFDCFSDST IRRAILGQVV RRRCEYPSGH NKVGSLQYLA
CRF01_AE_T   PDLADQLIHL HYFDCFSDSA IRRAILGQVV RRRCEYPSGH NKVGSLQYLA
CRF01_AE_T   PDLADKLIHL QYFGCFSDSA IRKAILGQVV RRRCEYPSGH NKVGSLQYLA
CRF01_AE_T   PDLADQLIHL QYFDCFSDSA IRKAILGQVV RRRCEYPSGH NKVGSLQYLA
CRF02_AG_F   PDLADQLIHL HYFDCFAESA IRKAILGEVV RPRCEYQAGH KQVGSLQYLA
CRF02_AG_F   PDLADQLIHL HYFDCFTDSA IRKAILGQVV SPRCEYQAGH NQVGSLQYLA
CRF02_AG_G   PDLADQLIHL HYFTCFSESA IRKAILGEVV RPRCEYQAGH NKVGSLQYLA
CRF02_AG_N   PDLADQLIHL YYFNCFSDSA IRKAILGEIV RPRCEYQAGH NKVGSLQYLA
CRF02_AG_S   PDLADQLIHL HYFDCFSDSA IRKAILGQIV RPRCEYQAGH TKVGSLQYLA
CRF02_AG_S   PDLADQLIHL HYFDCFSESA IRKALLGQVV RPKCEYQAGH NKVGSLQYLA
CRF03_AB_R   PNLADQLIHL YYFDCFSESA IRNAILGHRV SPSCEYRAGH NKVGSLQYLA
CRF03_AB_R   PNLADQLIHL YYFDCFSDSA IRNAILGHRV SPSCEYRAGH NKVGSLQYLA
CRF04_cpx_   PDLADQLIHM HYFDCFSESA IRKAILGHRV SPRCEYQAGH NKVGSLQYLA
CRF04_cpx_   PDLADQLIHM HYFDCFSESA IRQAILGYRV SPRCEYQAGH NEVGSLQYLA
CRF04_cpx_   PDLADQLIHV HYFDCFSESA IRKAILGHRV SPRCEYQAGH NKPGSLQYLA
CRF05_DF_B   PSLADQLIHV YYFDCFSESA IRNAILGRIV SPRCEFQAGH NKVGSLQYLA
CRF05_DF_B   PGLADQLIHM YYFDCFSESA IRKAILGYRV SPRCEYQAGH NKVGSLQYLA
CRF06_cpx_   PSMADQLIHI HYFDCFSESA IRKALLGHRV SPRCDYQAGH NKVGSLQYLA
CRF06_cpx_   PGLADQLIHM HYFDCFSESA IRKAILGQIV SPQCDYQAGH NKVGSLQYLA
CRF06_cpx_   PSLADQLIHM HYFDCFSESA IRKAILGHVV SPKCDYQAGH NKVGSLQYLA
CRF06_cpx_   PSLADQLIHM HYFDCFTESA IREAILGHIV SPRCDFEAGH NKVGSLQYLA
CRF11_cpx_   PGLADQLIHI HYFDCFSESA IREAILGHRV SPRCEYQAGH NQVGSLQYLA
CRF11_cpx_   PELADQLIHM HYFDCFAESA IRKAILGHRV SPRCEYPAGR NKVGSLQYLA
D_CD_84ZR0   PGLADQLIHM YYFDCFADSA IRKAILGHIV SPRCEYQAGH NKVGSLQYLA
D_CD_ELI_K   PGLADQLIHM YYFDCFSESA IRKAILGDIV SPRCEYQAGH NKVGSLQYLA
D_CD_NDK_M   PGLADQLIHM YYFDCFAESA IRKAILGHIV SPSCEYQAGH NKVGSLQYLA
D_UG_94UG1   PGLADQLIHI YHFDCFAESA IRKAILGQVV YPRCNYQAGH NKVGSLQYLA
F1_BE_VI85   PGLADQLIHI YYFDCFSESA IRKAILGHRI SPRCNYQAGH NKVGSLQYLA
F1_BR_93BR   PGLADQLIHI YYFDCFSESA IRKAILGHKI SPRCNYQAGH NKVGSLQYLA
F1_FI_FIN9   PGLADQLIHI YYFDCFSESA IRKAILGHRI SPRCDYQAGH NKVGSLQYLA
F1_FR_MP41   PDLADQLIHI YYFDCFSASA IRKAILGHRI SPRCNYQAGH NKVGSLQYLA
F2_CM_MP25   PGLADQLIHL HYFDCFSDSA IRKAILGQRV SPRCNYQAGH NKVGSLQYLA
F2KU_BE_VI   PGLADQLIHM HYFDCFTDSA IRKALIGLRV SPRCEYQAGH NKVGSLQYLA
G_BE_DRCBL   PDLADHLIHL HYFNCFSESA IRKAILGQTV RPSCEYPAGH NKVGSLQYLA
G_NG_92NG0   PNTADHLIHL YYFDCFSESA IRKAILGEIV SPRCEYPAGH NKVGSLQYLA
G_SE_SE616   PDLADHLIHL HYFDCFSDSA IRKAILGQIV SPRCEYQAGH NQVGSLQYLA
H_BE_VI991   PDLADQLIHL HYFDCFSDSA IRKAILGHRV SPICDYQAGH RKVGSLQYLA
H_BE_VI997   PGLADQLIHT HYFDCFSESA IRGAILGRVV SPRCEYQAGH NQVGSLQYLA
H_CF_90CF0   PGLADQLIHM HYFDCFSESA IRKAILGRVV RPRCNYPAGH KQVGTLQYLA
J_SE_SE702   PGLADQLIHM HYFDCFSDSA IRKAILGQIV SPRCDYQAGH NKVGSLQYLA
J_SE_SE788   PGLADQLIHM CYFDCFSDSA IRKAILGQIV SPRCDYQAGH NKVGSLQYLA
K_CD_EQTB1   PGLADQLIHI YYFDCFSESA IRKALLGHRV SPRCEYQAGH TQVGSLQYLA
```

```
K_CM_MP535   PDLADQLIHI YYFDCFSESA VRKAILGHRV SPRCECQAGH NKVGSLQYLA
N_CM_YBF30   PEMADKLIHL HYFDCFTASA IRQAVLGRPV LPRCEYPAGH KQVGTLQYLA
O_CM_ANT70   PETADRMIHL HYFTCFTASA VRKAILGQRV LTKCEYPTGH SQVGTLQLLA
O_CM_MVP51   PETADRMIHL HYFTCFTESA IRKAILGQRV LTKCEYLAGH SQVGTLQFLA
O_SN_99SE_   PETADRMIHI YYFACFTESA IRKAILGQRV LTRCEYPAGH SQVGTLQLLA
O_SN_99SE_   PETADRMIHT YYFACFTESA IRKAILGQRV LTRCEYSAGH SQVGTLQLLA
U_CD___83C   PDLADQLIHL HYFDCFSDSA IRKAILGHIV SPRCEYQTGH NKVGSLQYLA

             151                                                194
00BW0762_1   LTALIKPKKR KPPLPSVRKL VEDRWNKPQK TRGRRGNHTM NGH.
00BW0768_2   LTALIKPKKI KPPLPSVRKL VEDRWNKPQR TKGRRGNHTM NGH.
00BW0874_2   LTALIKPHKR KPPLPSVRKL VEDRWNNPQK TKGRRGNHTM NGH.
00BW1471_2   LTALIKPKRI KPPLPSLQKL VEDKWNNPQK TRGHRGSHTM NGH.
00BW1616_2   LTALIKPKKI KPPLPSVRKL VEDRWNNPQK TRGRRGNHTM NGH.
00BW1686_8   LTALIKPKKI KPPLPSIRKL VEDRWNKPQK TRGRRGNHTM NGH.
00BW1759_3   LTALIKPRKI KPPLPSVRKL VEDKWNKPQK TRGRRGNHTM NGH.
00BW1773_2   LTALIKPKKI KPPLPSVRKL VEDRWNNPQK TRGRRGNHTM NGH.
00BW1783_5   LTALIKPKKR KPPLPSVRKL VEDRWNKPPK TRDRRGNHTM NGH.
00BW1795_6   LTALIKPKKR KPPLPSVKKL VEDRWNKPQK TRGRRGSHTM NEH.
00BW1811_3   LTALIKPQRR KPPLPSVSKL VEDRWNNPQK TRGRRGCHTM NGH.
00BW1859_5   LTALIKPKKI KPPLPSVRKL VEDRWNNPQK TRGRRGNHTM NGH.
00BW1880_2   LTALIKPKKI KPPLPSVRKL VEDRWNKPQK TRGRRGNYTM NGH.
00BW1921_1   LTALIKPKKI KPPLPSVQKL VEDRWNKPQK TRGRRGNHTM NGH.
00BW2036_1   LTALIKPKKR KPPLPSVRKL VEDRWNKPQK TRGRKGNHTM NGH.
00BW2063_6   LTALIKPKKR KPPLPSVRKL VEDRWNNPQK TRGHRGNHTM NGH.
00BW2087_2   LTALVKPKKI KPPLPSVKKL VEDRWNKPQK TRGRRGNHTM NGR.
00BW2127_2   LTALIKPKQI KPPLPSVQKL VEDRWNKPQK TRGRRGDHTM NGH.
00BW2128_3   LTALIKPKKI KPPLPSVKKL VEDRWNNPQK TRGRRGNHTM SGH.
00BW2276_7   LTALIKPKRR KPPLPSVRKL VEDRWNKPQK TRGRRGNHTM NGH.
00BW3819_3   LTAIK.PKKR KPPLPSVRKL VEDRWNKPQK TRGRRGNHTM NGH.
00BW3842_8   LTALIKPKKR KPPLPSVRKL VEDRWNKSQK TRDRRGNHTM NGH.
00BW3871_3   LTALIKPKKI KPPLPSIRKL VEDRWNKSQK TRGRRGNHTM NGH.
00BW3876_9   LTALIKPKKI KPPLPSVRKL AEDRWNNPQK TRGRRGNHTM SGH.
00BW3886_8   LTALIKPKKR KPPLPSVRKL VEDRWNNSQK TRDHRGNHTM SGH.
00BW3891_6   LTALIKPKKR KPPLPSVRKL VEDRWNNPQK TRGHRGNHTM NGH.
00BW3970_2   LTTLIKPKRR KPPLPSVRKL AEDRWNNPQK TRDRRGNHTM NGH.
00BW5031_1   LTALIKPKRP KPPLPSVRKL AEDRWNKPRK TRGRRGNHTM NGH.
 96BW01B21   LTALIKPKKR KPPLPSVKKL VEDRWNDPQK TRGRRGSHTM NGH.
  96BW0407   LTALIKPKKR KPPLPSVRKL VEDRWNEPQK TRGRRGNHTM NGH.
  96BW0502   LTALIKPKQR KPPLPSVRKL VEDRWNKPQK TRGRRGNHTM NGH.
 96BW06_J4   LTALIKPKKR KPPLPSISKL VEDRWNKPQR TRGRRGNHTM NGH.
 96BW11_06   LTALVKPKKI KPPLPSVRKL VEDRWNKPQK TRGRRGNHTM NGH.
  96BW1210   LTALIKPKKR KPPLPSVRKL VEDRWNKPQK TRGRKGNHTM NGH.
 96BW15B03   LTALIKPKQI KPPLPSVRKL VEDRRNKPQK TRGRRGNRTM NGH.
 96BW16_26   LTALIKPKKI KPPLPSVNKL VEDRWNNPQK TRGRRGNHTL NGH.
 96BW17A09   LTAVIKPKKI KPPLPSVQKL VEDRWNKPQK TRGHRGSHTM NGH.
 96BWMO1_5   LTALIKPKKR KPPLPSVRKL VEDRWNKPQK TRGRRESHTM NGH.
 96BWMO3_2   LTALIKPKRI KPPLPSVRKL TEDRWNKPQK TKGRRGNHTM NGH.
98BWMC12_2   LTALIKPQKR KPPLPSVRKL VEDRWNNPQK TRGRRGNHTM NGH.
98BWMC13_4   LTALIKPKKR KPPLPSVKKL VEDRWNKPQK TRGRRGSHIM NGH.
98BWMC14_a   LTALIKTKKR KPPLPSVSKL VEDRWNKPQK TRGRRENHTM NGH.
98BWMO14_1   LTALIKPKKR KPPLPSVRKL VEDRWNKPQK TRGHRGNHTM NGH.
98BWMO18_d   LTALIKPKKI KPPLPSVKKL VEDRWNKPQK TRDRRGNHTM NGH.
98BWMO36_a   LTALIKPKRR KPPLPSVRKL VEDRWNKPQK TRGRRGNHTM NGH.
98BWMO37_d   LTALIKPKRR KPPLPSVRKL TEDRWNKPQK TRDHRGNHTM NGH.
99BW3932_1   LTALIKPKKI KPPLPSVQKL VEDRWNKPQK TRGRRGNHTM NGH.
99BW4642_4   LTALIKPKKI KPPLPSIRKL VEDRWNNPQK TRGRRGNHTM NGH.
99BW4745_8   LTALLKTKRR KPPLPSVRKL VEDRWNNPQK TRGHRGNHTM NGH.
99BW4754_7   LTALIKPKRI KPPLPSVRKL VEDRWNKPQK TRGHRGNHTM NGH.
99BWMC16_8   LTALIKPKVI KPPLPSVRKL VEDRWNKPQK TRGRRGNHTM NGH.
A2_CD_97CD   LRALVASTRT KPPLPSVRKL VEDRWNKPQK TRGHRGSHTM NGC.
```

```
A2_CY_94CY   LKAVVASTRT KPPLPSVRKL VEDRWNKPQK TKGHRGSHTM NGC.
A2D___97KR   LKALVGETRT KPPLPSVRKL TEDRWNKPQK TKGHRGSHTM NGH.
A2G_CD_97C   LRALVKPTKI KPPLPSVKKL TEDRWNKPQK TRGHRENPTM SGY.
A_BY_97BL0   LKALVTPTRE RPPLPSVRXL TEDRXNKPQK TRGRRXNHTM NXC.
A_KE_Q23_A   LKALVTPKKT KPPLPSVRIL TEDRWNKPQK TRGLRESHTM NGC.
A_SE_SE659   LRALVAPRKT KPPLPSVRIL AEDRWNKPQK TRDPRESHTM NGC.
A_SE_SE725   LKALVTPTRT KPPLPSVRKL AEDRWNKPQK TRGHRGSHTM NGC.
A_SE_SE753   LKALVTPKKT RPPLPSVRIL AEDRWNKSRK TRGPRGSHTM NGC.
A_SE_SE853   LKALVTPKKI KPPLPSVKKL TEDRWNKPQK TRGHRGNHTM HGY.
A_SE_SE889   LKALVTPKKI RPPLPSVRKL AEDRWNKPQK TKGHRGSHTM NGH.
A_SE_UGSE8   LKALVTPKRT KPPLPSVRKL TEDRWNKPQK TKGHRGSHTM NGC.
A_UG_92UG0   LKALVTPSRM KPPLPSVKKL AEDRWNKPQK TRGRRESHTM NGC.
A_UG_U455_   LKALVTPTRA KPPLPSVKKL TEDRWNKPQK TRGHRGSRTL NRH.
AC_IN_2130   LTALIKPKKR KPPLPSIRKL VEDRWNNPQK TRGRRGNHTM NGH.
AC_RW_92RW   LTALIKPKKI KPPLPSVSKL VEDKWNKPQK TRGRRGNHTM NGH.
AC_SE_SE94   LTALIKPKKI KPPLPSVRKL VEDKWNKPQK TRGRRGNHTM NGH.
ACD_SE_SE8   LKALVTPTRV KPPLPSVRKL AEDRWSKSQK TRGLRGSLTM NGC.
ACG_BE_VI1   LKALVTPTQI RPPLPSVRKL TEDRWNKPQK TRGHRGNHTM NGH.
AD_SE_SE69   LTALITPKKE KPPLPSVKKL TEDRWNKPQR TKGHRGSHTM NGH.
AD_SE_SE71   LKALVTPTKT KPPLPSVRIL TEDRWNKPQK TRGLRESHTM NGC.
ADHK_NO_97   LTALVAPKKI KPPLPSIKKL AEDRWNKPQK TRGHRGSHTM NGC.
ADK_CD_MAL   LTALIAPKKT RPPLPSVRKL TEDRWNKPQQ TKGHRGSHTM NGH.
AG_BE_VI11   LKALVTPTRI RPPLPSVRKL TEDRWNKPQK TRGHRGSHTM NGQW
AG_NG_92NG   LKALVTPTQT KPPLPSVKKL TEDRWNEPQK TRGHRGSHST NGH.
AGHU_GA_VI   LKALVTPTRE RPPLPSVQKL TEDRWNKPQK TKDHRGSHTM NGC.
AGU_CD_Z32   LTALITPKKT KPPLPSVKKL VEDRWNKPQK TRGHRENQTM NEH.
AJ_BW_BW21   LKAILKTEKR KPPLPSVQKL VEDRWNKPQR TRGHRESHTM NGH.
B_AU_VH_AF   LAALITPRQT KPPLPSVTKL TEDRWNKPRK TKGHRGSHTM SGH.
B_CN_RL42_   LTALTTPKNR KPPLPSVTKL TEDRWNKPQR TKGHRGSHTM SGH.
B_DE_D31_U   LAALITPKKI KPPLPSVAKL TEDRWNKPRK TKGHRGSHTM NGH.
B_DE_HAN_U   LAALTTPKKI KPPLPIVTKL TEDRWNKPQK TKGHRGSHTM HGH.
B_FR_HXB2_   LAALITPKKI KPPLPSVTKL TEDRWNKPQK TKGHRGSHTM NGH.
B_GA_OYI__   LAALIKPKKI KPPLPSVTKL TEDRWNKPQK TKGHRGSHTM NGH.
B_GB_CAM1_   LTALIAPKKI KPPLPSVRKL TEDRWNKPQK TKGHRGSHTM NGH.
B_GB_GB8_A   LTALITPKKI KPPLPSVTKL TEDRWNKPQK TKGHRGSHTM NGH.
B_GB_MANC_   LAALITPKKT KPPLPSVTKL TEDRWNKPQK TKGHRESHTM NGH.
B_KR_WK_AF   LTALITPKKI KPPLPSVRKL TEDRWNKPQK TKGHRGSHTM NGH.
B_NL_3202A   LAALIKPKKI KPPLPSVTKL TEDRWNKPQK TKGHRGSHTM NGH.
B_TW_TWCYS   LTALVQPKKI KPPLPSVVKL TEDRWNKPQK TKGHRGSHTM HGH.
B_US_BC_L0   LAALITPKRI KPPLPSVTKL TEDRWNKPQK TKGHRGSHTM NGH.
B_US_DH123   LAALVTPRKI KPPLPSVAKL TEDRWNKSHK TKGHRGSHTM NGH.
B_US_JRCSF   LTALIKPKKI KPPLPSVKKL TEDRWNKPQK TKGHRGSHTM NGH.
B_US_MNCG_   LTALITPKKI KPPLPSVKKL TEDRWNKPQK TKGHRGSHTI NGH.
B_US_P896_   LAALTTPRRI KPPFPSVTKL TEDRWNKPQK TKGHRGSHTM TGH.
B_US_RF_M1   LAALTTPKKI KPPLPSVKKL TEDRWNKPQK TKGHRGSHTM NGH.
B_US_SF2_K   LAALITPKKT KPPLPSVKKL TEDRWNKPQK TKGHRGSHTM NGH.
B_US_WEAU1   LTALITPKKI KPPLPSVKKL TEDRWNKPQK TKGHRGSHTM NGH.
B_US_WR27_   LTALIKPXKI KPPLPSVKKL TEDRWNXPQK TKGHRGSHTM NGH.
B_US_YU2_M   LTALITPKKT KPPLPSVKKL TEDRWNKPQK TKGHRGSRTM NGH.
BF1_BR_93B   LTALIKPKKR KPPLPSVKKL TEDRWNKPQK TKDHRGSHTM NGH.
C_BR_92BR0   LTALIKPKKI KPPLPSVKKL VEDRWNKPQK TRDRRGNHTM NGH.
C_BW_96BW0   LTALIKPKKR KPPLPSVRKL VEDRWNEPQK TRGRRGNHTM NGH.
C_BW_96BW1   LTALIKPKKI KPPLPSVRKL VEDRWNKPQK TRGRRGNHTM NGH.
C_BW_96BW1   LTALIKPKKR KPPLPSVRKL VEDRWNKPQK TRGRKGNHTM NGH.
C_BW_96BW1   LTALIKPKQI KPPLPSVRKL VEDRRNKPQK TRGRRGNRTM NGH.
C_ET_ETH22   LTALIKPKKA KPPLPSVSKL VEDKWNKPQK TRGRRGNHTM NGH.
C_IN_93IN1   LTALIKPKKI KPPLPSIKKL VEDRWNNPQK IRGRRGNHTM NGH.
C_IN_93IN9   LTALIKPKKI KPPLPSIKKL VEDRWNNPQK IRGRRGNHTM NGH.
C_IN_93IN9   LTALIKPKKI KPPLPSVRKL VEDRWNNPLK TRGRRGNHTM NGH.
C_IN_94IN1   LTALIKPKKI KPPLPSIKKL VEDRWNNPQK IRGRRGNHTM NGH.
C_IN_95IN2   LTALIKPKKI KPPLPSIKKL VEDRWNNPQK IRGRKGNHIM HGH.
```

```
CRF01_AE_C  LKALATPKKT RPPLPSVRKL TEDRWNKPQK TRGHRENPTM NGH.
CRF01_AE_C  LKALTKTKKT KPPLPSVRKL TEDRWNKPQK TKGHRESPTM NGH.
CRF01_AE_C  LKALATPKKI RPPLPSVRKL TEDRWNKPQK TRGHRENPTM SGH.
CRF01_AE_T  LKALTTPKRI RPPLPSVKKL TEDRWNKPQK IWGHRENPTM NGH.
CRF01_AE_T  LKALTTPKRI KPPLPSVKKL TEDRWNKPQK IRDHREYRTM NGH.
CRF01_AE_T  LKALTTPKRI RPPLPSVKKL TEDRWNKPQK IKGHRENPTM NGH.
CRF01_AE_T  LKALTTPKRI RPPLPSVKKL TEDRWNKHQ. KGDHRENPTM NGH.
CRF01_AE_T  LKALTTPKRI RPPLPSV.EI TEDRWNKPQ. KRGHRENPTM NGH.
CRF01_AE_T  LKALTTPKRI KPPLPSVRKL TEDRWNEPQK IRGHREYPTM NGH.
CRF02_AG_F  LKALVTPAKT KPPLPSVKKL AEDRWNKPQK TRGHRGNRSM NGH.
CRF02_AG_F  LKALVTPVKT KPPLPSVKKL AEDRWNKPQK TRGHRGNRSM NGQ.
CRF02_AG_G  LKALVTPTRK KPPLPSVRKL AEDRWNEPQK TRGHRGSRPM NGR.
CRF02_AG_N  LNALVAPTKT KPPLPSVRKL AEDRWKEPQK TRGHRGSRPM NGH.
CRF02_AG_S  LKALVTPTRT KPPLPSVKKL AEDRWNEPQK TRGHRGSRSM NGH.
CRF02_AG_S  LKALVTPTRR KPPLPSVKKL AEDRWNEPQK TRGHRGNRSM NGH.
CRF03_AB_R  LAALRTPKKI KPPLPSVTKL TEDRWNKPQR TKDHRGSHTM SGH.
CRF03_AB_R  LAALRTPKKI KPPLPSVTKL TEDRWNKPQR TKDHRGSHTM SGH.
CRF04_cpx_  LAALISPKKT KPPLPSVKKL VEDRWNKPQK TRGRRENQIM NGH.
CRF04_cpx_  LAALISPKKT KPPLPSVKKL VEDRWNKPQK TRGRRENQIM NGH.
CRF04_cpx_  LAALISPKKT KPPLPSVKKL VEDRWNKSQK TKGRRESHIM NGH.
CRF05_DF_B  LTALITPKKT KPPLPSVRKL TEDRWNKPQK TKGRRGNHTM NGY.
CRF05_DF_B  LTALITPQKI KPPLPSVRKL TEDRWNKPQR TKGHRGCHTM NGY.
CRF06_cpx_  LTALIKPEKR KPPLPSVQKL VEDRWNKPQK TRGHRESHTM NGH.
CRF06_cpx_  LTALIKPKKR KPPLPSVQKL VEDRWNKPQK TRDHRESHTM NGH.
CRF06_cpx_  LTALIKPRKR KPPLPSVQKL VEDRWNKPQK TRDHRECHTM NGH.
CRF06_cpx_  LKALVKTKRR KPPLPSVQKL VEDRWNKPQK TKDHRESHIM DGH.
CRF11_cpx_  LKALVTPTRA KPPLPSVRKL AEDRWNKPQK TRGHRGNHTA NGC.
CRF11_cpx_  LKALVTPKRT KPPLPSVRKL TEDRWNKPQK TRGRRGNHTV NGC.
D_CD_84ZR0  LTALIAPKKR KPPLPSVKKL TEDRWNKPRQ TKGRRGSHTM NGH.
D_CD_ELI_K  LTALIAPKQI KPPLPSVRKL TEDRWNKPQQ TRGHRGSHTM NGH.
D_CD_NDK_M  LAALIAPKKI KPPLPSVRKL TEDRWNKPQK TKGRRGSHTM NGH.
D_UG_94UG1  LTALVTPRKI KPPLPSVGKL TEDRWNKPQR TKGHRGSHTM NGH.
F1_BE_VI85  LTALIAPEKT KPPLPSVQKL VEDRWNKPQE TRGHRGSHTM NGH.
F1_BR_93BR  LTALIAPKKT KPPLPSVQKL VEDRWNKPQK TRGHRESHTM NGH.
F1_FI_FIN9  LTALVSPKKA KPPLPSVKKL VEDRWNKPQE IRGHRGSHTM NGH.
F1_FR_MP41  LTALIAPKKT KPPLPSVKKL VEDRWNKPQE TRGHRGSHTM NGH.
F2_CM_MP25  LTALITPKKI KPPLPSVRKL VEDRWNNPQK TRGHRGSHTM NGH.
F2KU_BE_VI  LTALVAPKKT KPPLLSVRKL VEDRWNKPQK TRDHRGSHTM NGH.
G_BE_DRCBL  LKVLVAPTRR RPPLPSVRKL TEDRWNEPQK TRGHRENPTM NGH.
G_NG_92NG0  SKALVTPTRK RPPLPSVGKL AEDRWNKPQK TRDHRENPTM NGH.
G_SE_SE616  LKVLVTSKRS RPPLPSVTEL AEDRWNKPQK TRGHRENPTM NGH.
H_BE_VI991  LTALISPKRT KPPLPSVRKL VEDRWNKPQK TRGHRGSHTM NGH.
H_BE_VI997  LTALVAPKKT KPPLPSVKKL VEDGWNKPQK TRGHRGSHTM NRH.
H_CF_90CF0  LTALVAPKKI KPPLPSVRKL VEDRWNKPQK TRGHRGSHTM NGH.
J_SE_SE702  LTALIKPKRR KPPLPSVQKL VEDRWNKPQK TRDHRESHTM NGH.
J_SE_SE788  LTALIRPKRR KPPLPSVQKL VEDRWNKPQK TTGHRESHTM NGH.
K_CD_EQTB1  LTALIAPKKT KPPVPSVQKL VEDRWNKPQK TRGHRGSHTM SGQ.
K_CM_MP535  LTALVAPRRP KPPVPSVKKL VEDRWNKPQK TRGHRGSQTM NGH.
N_CM_YBF30  LTAWVGAKKR KPPLPSVTKL TEDRWNEHQK MQGHRGNPIM NGH.
O_CM_ANT70  LRAVVKARSR KPPLPSVQKL TEDRWNKHLR IRDQLKSPSM NGH.
O_CM_MVP51  LKAVVKVKRN KPPLPSVQRL TEDRWNKPWK IRDQLGSHSM NGH.
O_SN_99SE_  LRVVVKEKRN KPPLPSVQKL TEDRWSRHLR IRDQLESHSM NGH.
O_SN_99SE_  LRVVVKEKRH KPPLPSVQKL TEDRWSRHLR IRDQLGSHSM NGH.
U_CD___83C  LTTLVAPTKR KPPLPSVRKL VEDRWNKPQK TKGHKGSHTM HGH.
```

Table 18. HIV Vpr Sequence Alignment
GCG Multiple Sequence File.
Written by Omiga 1.1

Name: 00BW0762_1	Len:	100	Check:	8179	Weight:	1.00
Name: 00BW0768_2	Len:	100	Check:	8119	Weight:	1.00
Name: 00BW0874_2	Len:	100	Check:	7661	Weight:	1.00
Name: 00BW1471_2	Len:	100	Check:	6614	Weight:	1.00
Name: 00BW1616_2	Len:	100	Check:	6361	Weight:	1.00
Name: 00BW1686_8	Len:	100	Check:	6014	Weight:	1.00
Name: 00BW1759_3	Len:	100	Check:	6894	Weight:	1.00
Name: 00BW1773_2	Len:	100	Check:	7772	Weight:	1.00
Name: 00BW1783_5	Len:	100	Check:	7149	Weight:	1.00
Name: 00BW1795_6	Len:	100	Check:	7614	Weight:	1.00
Name: 00BW1811_3	Len:	100	Check:	7968	Weight:	1.00
Name: 00BW1859_5	Len:	100	Check:	6222	Weight:	1.00
Name: 00BW1880_2	Len:	100	Check:	6941	Weight:	1.00
Name: 00BW1921_1	Len:	100	Check:	8183	Weight:	1.00
Name: 00BW2036_1	Len:	100	Check:	8175	Weight:	1.00
Name: 00BW2063_6	Len:	100	Check:	8705	Weight:	1.00
Name: 00BW2087_2	Len:	100	Check:	7388	Weight:	1.00
Name: 00BW2127_2	Len:	100	Check:	8282	Weight:	1.00
Name: 00BW2128_3	Len:	100	Check:	1723	Weight:	1.00
Name: 00BW2276_7	Len:	100	Check:	6468	Weight:	1.00
Name: 00BW3819_3	Len:	100	Check:	5670	Weight:	1.00
Name: 00BW3842_8	Len:	100	Check:	7788	Weight:	1.00
Name: 00BW3871_3	Len:	100	Check:	8574	Weight:	1.00
Name: 00BW3876_9	Len:	100	Check:	7285	Weight:	1.00
Name: 00BW3886_8	Len:	100	Check:	6446	Weight:	1.00
Name: 00BW3891_6	Len:	100	Check:	8629	Weight:	1.00
Name: 00BW3970_2	Len:	100	Check:	7113	Weight:	1.00
Name: 00BW5031_1	Len:	100	Check:	5511	Weight:	1.00
Name: 96BW01B21	Len:	100	Check:	7551	Weight:	1.00
Name: 96BW0407	Len:	100	Check:	8226	Weight:	1.00
Name: 96BW0502	Len:	100	Check:	8242	Weight:	1.00
Name: 96BW06_J4	Len:	100	Check:	7544	Weight:	1.00
Name: 96BW11_06	Len:	100	Check:	7942	Weight:	1.00
Name: 96BW1210	Len:	100	Check:	8580	Weight:	1.00
Name: 96BW15B03	Len:	100	Check:	7308	Weight:	1.00
Name: 96BW16_26	Len:	100	Check:	7009	Weight:	1.00
Name: 96BW17A09	Len:	100	Check:	6492	Weight:	1.00
Name: 96BWMO1_5	Len:	100	Check:	5837	Weight:	1.00
Name: 96BWMO3_2	Len:	100	Check:	5277	Weight:	1.00
Name: 98BWMC12_2	Len:	100	Check:	7807	Weight:	1.00
Name: 98BWMC13_4	Len:	100	Check:	9051	Weight:	1.00
Name: 98BWMC14_a	Len:	100	Check:	7867	Weight:	1.00
Name: 98BWMO14_1	Len:	100	Check:	7266	Weight:	1.00
Name: 98BWMO18_d	Len:	100	Check:	7638	Weight:	1.00
Name: 98BWMO36_a	Len:	100	Check:	7495	Weight:	1.00
Name: 98BWMO37_d	Len:	100	Check:	6640	Weight:	1.00
Name: 99BW3932_1	Len:	100	Check:	6974	Weight:	1.00
Name: 99BW4642_4	Len:	100	Check:	6081	Weight:	1.00
Name: 99BW4745_8	Len:	100	Check:	8860	Weight:	1.00
Name: 99BW4754_7	Len:	100	Check:	6856	Weight:	1.00
Name: 99BWMC16_8	Len:	100	Check:	8223	Weight:	1.00

```
                1                                                   50
00BW0762_1      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH SLGQHIYNTY
00BW0768_2      MEQAPEDQGP QREPYNEWTL EILEELKQEA VRHFPRPWLH NLGEYIYETY
00BW0874_2      MEQPPEDQGP QREPYNEWTL EILEELKQEA VRHFPRPWLH SLGQYIYETY
00BW1471_2      MEQPPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH SLGQHIYETY
00BW1616_2      MEQPPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH SLGQYIYENY
00BW1686_8      MEQAPEDQGP QREPYNEWAL EILEELKQEA VRHFPRPWLH SIGQYIYETY
00BW1759_3      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH GLGQHIYETY
00BW1773_2      MEQPPEDQGP QREPYNEWTL ELLEELIQEA VRHFPRPWLH SLGQYIYETY
00BW1783_5      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH SMGQHIYNTY
00BW1795_6      MEQAPEDQGP QREPYN.ETL ELLEELKQEA VRHFPRIWLH NLGQYIYNTY
00BW1811_3      MEQPPEDQGP QRVPYNEWAL ELLEELKQEA VRHFPRPWLH GLGQYVYETY
00BW1859_5      MEQPPEDQGP QREPYNEWAL EILEELKQEA VRHFPRLWLH SLGQYIYETY
00BW1880_2      MEQAPEDQGP QRELYNEWTL ELLEELKQEA ARHFPSSWLH GLGQHIYNTY
00BW1921_1      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRTWLH NLGQYIYQTY
00BW2036_1      MEQAPEDQGP QREPYNEWTL EILEELKQEA VRHFPRPWLQ SLGQYIYETY
00BW2063_6      MEQPPEDQGP QREPYNEWTL GLLEELKQEA VRHFPRLWLH NLGQYIYNTY
00BW2087_2      MEQAPEDQGP QREPYNEWAL ELLEELKQEA VRHFPRPWLH NLGQYIYETY
00BW2127_2      MEQAPEDQGP QRGPYNEWTL EILEELKQEA VRHFPRPWLH NLGQYIYETY
00BW2128_3      MEQPPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH GLGQYIYETY
00BW2276_7      MEQTPEDQGP QREPYNEWAL EILEELKQEA VRHFPRTWLH SLGQYIYDTY
00BW3819_3      MEQAPEDQGP QREPYNEWTL EILEELKQGA VRHFPRPWLH NLGQHIYETY
00BW3842_8      MEQVPEDQGP QREPYNEWTL EILEELKQEA VRHFPRPWLQ GLGHYIYETY
00BW3871_3      MEQVPEDQGP QREPYNEWTL EILEELKQEA VRHFPRPWLH NLGQYIYETY
00BW3876_9      MEQSPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH GIGQYIYETY
00BW3886_8      MEQFPEDQGP QREPYNEWTL ELLEELKQEA VKHFPRPWLH NLGQHIYETY
00BW3891_6      MEQPPEDQGP QREPYNEWTL EVLEELKQEA VRHFPRPWLH SLGQYVYETY
00BW3970_2      MEQPPEDQGP QREPYNEWAL EILEELKQEA VRHFPRPWLH SLGQHIYETY
00BW5031_1      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH SLGQHIYETY
 96BW01B21      MERPPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH GLGQYIYETY
  96BW0407      MERAPEDQGP QREPYNEWAL ELLEELKQEA VRHFPRMWLH GLGQYIYETY
  96BW0502      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPGPWLH GLGQYVYETY
 96BW06_J4      MEQAPEDQGP QREPYNEWTL EILEELKQEA VRHFPPPWLH SLGQYIYETY
 96BW11_06      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH SLGQHIYNTY
  96BW1210      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH SLGQYIYETY
 96BW15B03      TEQAPEDQGP QREPYNEWAL EILEELKQEA VRHFPRPWLH SLGQYIYETY
 96BW16_26      MEQPPEDQGP QREPYTEWAL ELLEELKQEA VRHFPRPWLH GLGQYIYDTY
 96BW17A09      MEQTPEDQGP QREPHNEWTL ELLEELKQEA VRHFPRPWLH SLGQHIYETY
 96BWMO1_5      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPR.TLH DLGQHIYNTY
 96BWMO3_2      MEQAPEDQGP QREPYNEWTL EILEELKQEA IRHFPIPYLQ HLGQYIYETY
98BWMC12_2      MEQPPEDQGP QREPYNEWTL EILEELKQEA VRHFPRPWLH SLGQYIYETY
98BWMC13_4      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRIWLH NLGQYVYNTY
98BWMC14_a      MEQAPEDQGP QREPYNEWTL EILEELKQEA VRHLPRPWLH SLGQHIYETY
98BWMO14_1      MEQAPEDQGP QREPYNEWTL ALLEDLKQEA VRHVPRPWLH SLGQHIYETY
98BWMO18_d      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH SLGQYIYETY
98BWMO36_a      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPITWLH NLGQYIYETY
98BWMO37_d      MEQAPEDQGP QREPYNEWTL EILEELKQEA VRHFLRPWLH DLGQYIYETY
99BW3932_1      MEQAPEDQGP QREPYNEWTL EILEELKQEA VRHFPRPWLH NLGQYIYATY
99BW4642_4      MEQPPEDQGP QREPYNEWAL EILEELKQEA VRHFPRPWLH NLGQYIYETY
99BW4745_8      MEQPPEDQGP QREPYNEWTL EVLEDLKQEA VRHFPRPWLH SIGQYVYSTY
99BW4754_7      MEQAPENQGP QREPYNEWAL ELLEELKQEA VRHFPRPWLH DLGQHIYNTY
99BWMC16_8      MEQAPEDQGP QREPYNEWTL ELLEELKQEA VRHFPRPWLH SLGLYIYETY

                51                                                  100
00BW0762_1      GDTWTGVEAI IRILQQLLFI HFRIGCQHSR IGIMRQ.... RRTRNGASRS
00BW0768_2      GDTWTGVEAL IRVLQQLLFI HFRIGCSHSR IGIVRQ.... RRARNGSSRS
00BW0874_2      GDTWTGVETI IRTLQQLLFI HFRIGCQHSR IGILRQ.... KRARNGASRS
00BW1471_2      GDTWAGVEAL LRILQQLLFI HFRIGCQHSR IGIIPQ.... RRARNGSRRS
00BW1616_2      GDTWAGVEAI TRILQQLLFI HFRIGCQHSR IGILRQ.... RRARNGANRS
00BW1686_8      GDTWTGVEAL MRILQQLLFI HFRIGCQHSR IGILQR.... R.ARNGASRS
00BW1759_3      GDTWTGVEAI IRILQQLLFI HYRIGCQHSR IGIVRQ.... RRARNGANRS
```

```
00BW1773_2  GDTWTGVEAI IKILQQLLFI HFRIGCQHSR IGILRQ.... RRARNGASRS
00BW1783_5  GDTWAGVEAI IRILQQLLFI HFRIGCQHSR IGILRQ.... RRTRNGASRS
00BW1795_6  GDTWTGVEAI IRTLQQLLFV HFRIGCQHSR IGIMRQ.... RRARNGTSGS
00BW1811_3  GDTWTGVEAI IRILQQLLFV HFRIGCQHSR IGILQQ.... RRARNGASRS
00BW1859_5  GDTWAGVEAL IRILQQLLFI HFRIGCQHSR IGILQQ.... RRARNGASRS
00BW1880_2  GDTWTGVEVL IRILQQLLFI HFRIGCQHSR IGIIRQ.... RRTRNGASRP
00BW1921_1  GDTWTGVEAL IRILQQPLFI HFRIGCQHSR IGITLP.... RRARNGANRS
00BW2036_1  GDTWTGVEAI IRILQQLLFI HFRSGCAHSR IGTLPQ.... RRARNGASRS
00BW2063_6  GDTWTGVEAI IRILQQLLFI HFRIGCQHSR IGIIRQ.... RRTRNGDSRS
00BW2087_2  GDTWTGVEAL IRILQQLLFT HYRFGCQHSR IGILQQ.... RRARNGANRS
00BW2127_2  GDTWTGVEVI IRILQQLLFI HFRIGCQHSR IGILRQ.... RRTRNGASRS
00BW2128_3  GDTWAGVESL IRMLQHLLFI HFRIGCQHSR IDX..... ..........
00BW2276_7  GDTWAGVEAI IRILQQLLFT HFRIGCHHSR IGILRQ.... RRARNGASRS
00BW3819_3  GDTWAGVEAL LRILQQLLFI HFRIGCQHSR IGILRQ.... RRARNGASRP
00BW3842_8  GDTWTGVETI IRILQQLLFI HFRIGCSRSR IGPMRQ.... RRARNGASRS
00BW3871_3  GDTWTGVEAL LRVLQQLLFV HFRIGCQHSR IGILQQ.... RRARNGSSRS
00BW3876_9  GDTWTGVEAI IRILQQLLFI HYRIGCAHSR IGIVRQ.... RRARNGANRS
00BW3886_8  GDTWTGVEAI IRMLQQLLFI HFRIGCQHSR IGILRQ.... RRARNGANRS
00BW3891_6  GDTWTGVEAL IRMLQQLLFI HFKIGCQHSR IGILRR.... RRARNGASRS
00BW3970_2  GDTWTGVEAL IRILQQLLFI HFRIGCQHSR IGIILQ.... RRTRNGASRS
00BW5031_1  GDTWMGVEAL IRILQ..... HFRIGCQHSR IGIILQ.... RRTRNGASRS
 96BW01B21  GDTWTGVENM IRILQQLLFV HFRIGCQHSR IGILQQ.... RRARNGASRS
  96BW0407  GDTWTGVEAL IRTLQQLLFI HFRIGCQHSR IGILRQ.... RRVRNGTNRS
  96BW0502  GDTWTGVETL IRILQQLLFI HFRIGCQHSR IGILRQ.... RRTRNGASRS
 96BW06_J4  GDTWTGVETI IRILQQLLFI HFRIGCQHSR IGILQQ.... RRARNGASRP
 96BW11_06  GDTWTGVEAI IRILQQLLFI HFRIGCQHSR IGIIRQ.... RRTRNGASRP
  96BW1210  GDTWTGVEVL TRILQQLLFI HFRIGCQHSR IGILRQ.... RRTRNGASRS
 96BW15B03  GDTWTGVEAI IRILQQLLFI HFRIGCLHSR IGIMRQ.... RRARNGASRS
 96BW16_26  GDTWTGVEIK IRILQQLLFI HFRIGCQHSR IGILQQ.... RRARNGARRS
 96BW17A09  GDTWAGVEAL LRILQQLLFI HFRIGCHHSR IGITPQ.... RRARNGSRRS
 96BWMO1_5  GDTWTGVEAI TRILQQLLFI HYRIGCQHSR IGIMRQ.... RRARNGASRS
 96BWMO3_2  GDTWAGVLAI IRILQQLLFI HFRIGCSHSR IGIWR..... RRARNGASRS
98BWMC12_2  GDTWTGVEAI LRILQQLLFI HFRIGCQHSR IGILRQ.... RRARNGASRS
98BWMC13_4  GDTWTGVEAI IRILQQLLFI HFRIGCQHSR IGILRQ.... RRTRNGASRS
98BWMC14_a  GDTWTGVEAI IRILQQLLFI HFRIGCQHSR IGILPR.... RRARNGSSRS
98BWMO14_1  GDTWTGVEAI IRILQQLLFI HFRIGCQHSR IGILRQ.... RRARNGANRS
98BWMO18_d  GDTWTGVEVI IRILQQLLFI HFRIGCQHSR IGILRQ.... RRARNGANRS
98BWMO36_a  GDTWTGVEAL IRTLQQLLFI HFRIGCQHSR IGILRQ.... RRARNGASRS
98BWMO37_d  GDTWTGVETI IRVLQQLLFI HFRIGCH.SR IGIVRQ.... RRARNGASRS
99BW3932_1  GDTWMGVEAL LRILQQLLFI HFRIGCQHSR IGILRQ.... RRARNGASRS
99BW4642_4  GDTWAGVEAI IRVLQQLLFI HFRIGCHHSR IGIMQQ.... RRARNGASRS
99BW4745_8  GDTWTGVEAL MRILQQLLFI HFRIGCRHSR IGILRQ.... RGARNGASRS
99BW4754_7  GDTWTGVEAI IRILQQLLFI HFRIGCHHSR IGIIRQ.... RRTRNGASRP
99BWMC16_8  GDTWTGVEVI IRILQQLLFI HFRIGCQHSR IGILRQ.... RRARNGPSRS
```

Table 19. HIV Vpu Sequence Alignment
GCG Multiple Sequence File.
Written by Omiga 1.1

Name:	00BW0762_1	Len:	106	Check:	4772	Weight:	1.00
Name:	00BW0768_2	Len:	106	Check:	7115	Weight:	1.00
Name:	00BW0874_2	Len:	106	Check:	7209	Weight:	1.00
Name:	00BW1471_2	Len:	106	Check:	7341	Weight:	1.00
Name:	00BW1616_2	Len:	106	Check:	3870	Weight:	1.00
Name:	00BW1686_8	Len:	106	Check:	8787	Weight:	1.00
Name:	00BW1759_3	Len:	106	Check:	7584	Weight:	1.00
Name:	00BW1773_2	Len:	106	Check:	7507	Weight:	1.00
Name:	00BW1783_5	Len:	106	Check:	7874	Weight:	1.00
Name:	00BW1795_6	Len:	106	Check:	8721	Weight:	1.00
Name:	00BW1811_3	Len:	106	Check:	3657	Weight:	1.00
Name:	00BW1859_5	Len:	106	Check:	6137	Weight:	1.00
Name:	00BW1880_2	Len:	106	Check:	5827	Weight:	1.00
Name:	00BW1921_1	Len:	106	Check:	4605	Weight:	1.00
Name:	00BW2036_1	Len:	106	Check:	6145	Weight:	1.00
Name:	00BW2063_6	Len:	106	Check:	7187	Weight:	1.00
Name:	00BW2087_2	Len:	106	Check:	9545	Weight:	1.00
Name:	00BW2127_2	Len:	106	Check:	4898	Weight:	1.00
Name:	00BW2276_7	Len:	106	Check:	7311	Weight:	1.00
Name:	00BW3819_3	Len:	106	Check:	4879	Weight:	1.00
Name:	00BW3842_8	Len:	106	Check:	1804	Weight:	1.00
Name:	00BW3871_3	Len:	106	Check:	6650	Weight:	1.00
Name:	00BW3876_9	Len:	106	Check:	6684	Weight:	1.00
Name:	00BW3886_8	Len:	106	Check:	8701	Weight:	1.00
Name:	00BW3891_6	Len:	106	Check:	8544	Weight:	1.00
Name:	00BW3970_2	Len:	106	Check:	9375	Weight:	1.00
Name:	00BW5031_1	Len:	106	Check:	7778	Weight:	1.00
Name:	96BW01B21	Len:	106	Check:	6481	Weight:	1.00
Name:	96BW0407	Len:	106	Check:	4225	Weight:	1.00
Name:	96BW0502	Len:	106	Check:	5292	Weight:	1.00
Name:	96BW06_J4	Len:	106	Check:	5367	Weight:	1.00
Name:	96BW11_06	Len:	106	Check:	6477	Weight:	1.00
Name:	96BW1210	Len:	106	Check:	6400	Weight:	1.00
Name:	96BW15B03	Len:	106	Check:	2981	Weight:	1.00
Name:	96BW16_26	Len:	106	Check:	5352	Weight:	1.00
Name:	96BW17A09	Len:	106	Check:	6778	Weight:	1.00
Name:	96BWMO1_5	Len:	106	Check:	5954	Weight:	1.00
Name:	96BWMO3_2	Len:	106	Check:	6334	Weight:	1.00
Name:	98BWMC12_2	Len:	106	Check:	6905	Weight:	1.00
Name:	98BWMC13_4	Len:	106	Check:	7458	Weight:	1.00
Name:	98BWMC14_a	Len:	106	Check:	4023	Weight:	1.00
Name:	98BWMO14_1	Len:	106	Check:	5708	Weight:	1.00
Name:	98BWMO18_d	Len:	106	Check:	7741	Weight:	1.00
Name:	98BWMO36_a	Len:	106	Check:	5445	Weight:	1.00
Name:	98BWMO37_d	Len:	106	Check:	8225	Weight:	1.00
Name:	99BW3932_1	Len:	106	Check:	1112	Weight:	1.00
Name:	99BW4642_4	Len:	106	Check:	8891	Weight:	1.00
Name:	99BW4745_8	Len:	106	Check:	3424	Weight:	1.00
Name:	99BW4754_7	Len:	106	Check:	5468	Weight:	1.00
Name:	99BWMC16_8	Len:	106	Check:	6656	Weight:	1.00
Name:	A2_CD_97CD	Len:	106	Check:	6086	Weight:	1.00
Name:	A2_CY_94CY	Len:	106	Check:	4609	Weight:	1.00
Name:	A2D___97KR	Len:	106	Check:	2120	Weight:	1.00
Name:	A2G_CD_97C	Len:	106	Check:	4405	Weight:	1.00
Name:	A_BY_97BL0	Len:	106	Check:	913	Weight:	1.00
Name:	A_KE_Q23_A	Len:	106	Check:	3380	Weight:	1.00
Name:	A_SE_SE659	Len:	106	Check:	4192	Weight:	1.00

Name:	A_SE_SE725	Len:	106	Check:	3244	Weight: 1.00
Name:	A_SE_SE753	Len:	106	Check:	1918	Weight: 1.00
Name:	A_SE_SE853	Len:	106	Check:	5495	Weight: 1.00
Name:	A_SE_SE889	Len:	106	Check:	5422	Weight: 1.00
Name:	A_SE_UGSE8	Len:	106	Check:	4254	Weight: 1.00
Name:	A_UG_92UG0	Len:	106	Check:	4081	Weight: 1.00
Name:	A_UG_U455_	Len:	106	Check:	2987	Weight: 1.00
Name:	AC_IN_2130	Len:	106	Check:	7929	Weight: 1.00
Name:	AC_RW_92RW	Len:	106	Check:	5133	Weight: 1.00
Name:	AC_SE_SE94	Len:	106	Check:	7394	Weight: 1.00
Name:	ACD_SE_SE8	Len:	106	Check:	1852	Weight: 1.00
Name:	ACG_BE_VI1	Len:	106	Check:	6357	Weight: 1.00
Name:	AD_SE_SE69	Len:	106	Check:	5734	Weight: 1.00
Name:	AD_SE_SE71	Len:	106	Check:	4697	Weight: 1.00
Name:	ADHK_NO_97	Len:	106	Check:	6301	Weight: 1.00
Name:	ADK_CD_MAL	Len:	106	Check:	4338	Weight: 1.00
Name:	AG_BE_VI11	Len:	106	Check:	3500	Weight: 1.00
Name:	AG_NG_92NG	Len:	106	Check:	4859	Weight: 1.00
Name:	AGHU_GA_VI	Len:	106	Check:	6173	Weight: 1.00
Name:	AGU_CD_Z32	Len:	106	Check:	9411	Weight: 1.00
Name:	AJ_BW_BW21	Len:	106	Check:	6158	Weight: 1.00
Name:	B_AU_VH_AF	Len:	106	Check:	5007	Weight: 1.00
Name:	B_CN_RL42_	Len:	106	Check:	4249	Weight: 1.00
Name:	B_DE_D31_U	Len:	106	Check:	4572	Weight: 1.00
Name:	B_DE_HAN_U	Len:	106	Check:	6819	Weight: 1.00
Name:	B_FR_HXB2_	Len:	106	Check:	5240	Weight: 1.00
Name:	B_GA_OYI_M	Len:	106	Check:	5651	Weight: 1.00
Name:	B_GB_CAM1_	Len:	106	Check:	5359	Weight: 1.00
Name:	B_GB_GB8_A	Len:	106	Check:	1955	Weight: 1.00
Name:	B_GB_MANC_	Len:	106	Check:	6521	Weight: 1.00
Name:	B_KR_WK_AF	Len:	106	Check:	2320	Weight: 1.00
Name:	B_NL_3202A	Len:	106	Check:	4510	Weight: 1.00
Name:	B_TW_TWCYS	Len:	106	Check:	5491	Weight: 1.00
Name:	B_US_BC_L0	Len:	106	Check:	3142	Weight: 1.00
Name:	B_US_DH123	Len:	106	Check:	4669	Weight: 1.00
Name:	B_US_JRCSF	Len:	106	Check:	4070	Weight: 1.00
Name:	B_US_MNCG_	Len:	106	Check:	3291	Weight: 1.00
Name:	B_US_P896_	Len:	106	Check:	2280	Weight: 1.00
Name:	B_US_RF_M1	Len:	106	Check:	3104	Weight: 1.00
Name:	B_US_SF2_K	Len:	106	Check:	3857	Weight: 1.00
Name:	B_US_WEAU1	Len:	106	Check:	3817	Weight: 1.00
Name:	B_US_WR27_	Len:	106	Check:	3329	Weight: 1.00
Name:	B_US_YU2_M	Len:	106	Check:	5184	Weight: 1.00
Name:	BF1_BR_93B	Len:	106	Check:	3243	Weight: 1.00
Name:	C_BR_92BR0	Len:	106	Check:	7645	Weight: 1.00
Name:	C_BW_96BW0	Len:	106	Check:	5235	Weight: 1.00
Name:	C_BW_96BW1	Len:	106	Check:	6477	Weight: 1.00
Name:	C_BW_96BW1	Len:	106	Check:	6400	Weight: 1.00
Name:	C_BW_96BW1	Len:	106	Check:	2981	Weight: 1.00
Name:	C_ET_ETH22	Len:	106	Check:	8303	Weight: 1.00
Name:	C_IN_93IN1	Len:	106	Check:	8376	Weight: 1.00
Name:	C_IN_93IN9	Len:	106	Check:	6231	Weight: 1.00
Name:	C_IN_93IN9	Len:	106	Check:	7626	Weight: 1.00
Name:	C_IN_94IN1	Len:	106	Check:	6889	Weight: 1.00
Name:	C_IN_95IN2	Len:	106	Check:	8199	Weight: 1.00
Name:	CRF01_AE_C	Len:	106	Check:	4437	Weight: 1.00
Name:	CRF01_AE_C	Len:	106	Check:	4082	Weight: 1.00
Name:	CRF01_AE_C	Len:	106	Check:	3725	Weight: 1.00
Name:	CRF01_AE_T	Len:	106	Check:	3201	Weight: 1.00
Name:	CRF01_AE_T	Len:	106	Check:	3137	Weight: 1.00
Name:	CRF01_AE_T	Len:	106	Check:	3484	Weight: 1.00

```
Name: CRF01_AE_T       Len:   106  Check: 3491  Weight:   1.00
Name: CRF01_AE_T       Len:   106  Check: 2300  Weight:   1.00
Name: CRF01_AE_T       Len:   106  Check: 2481  Weight:   1.00
Name: CRF02_AG_F       Len:   106  Check: 2748  Weight:   1.00
Name: CRF02_AG_F       Len:   106  Check: 4618  Weight:   1.00
Name: CRF02_AG_G       Len:   106  Check: 400   Weight:   1.00
Name: CRF02_AG_N       Len:   106  Check: 5979  Weight:   1.00
Name: CRF02_AG_S       Len:   106  Check: 5296  Weight:   1.00
Name: CRF02_AG_S       Len:   106  Check: 4213  Weight:   1.00
Name: CRF03_AB_R       Len:   106  Check: 952   Weight:   1.00
Name: CRF03_AB_R       Len:   106  Check: 431   Weight:   1.00
Name: CRF04_cpx_       Len:   106  Check: 6986  Weight:   1.00
Name: CRF04_cpx_       Len:   106  Check: 8606  Weight:   1.00
Name: CRF04_cpx_       Len:   106  Check: 5826  Weight:   1.00
Name: CRF05_DF_B       Len:   106  Check: 5193  Weight:   1.00
Name: CRF05_DF_B       Len:   106  Check: 5092  Weight:   1.00
Name: CRF06_cpx_       Len:   106  Check: 3214  Weight:   1.00
Name: CRF06_cpx_       Len:   106  Check: 3831  Weight:   1.00
Name: CRF06_cpx_       Len:   106  Check: 4862  Weight:   1.00
Name: CRF06_cpx_       Len:   106  Check: 3114  Weight:   1.00
Name: CRF11_cpx_       Len:   106  Check: 4492  Weight:   1.00
Name: CRF11_cpx_       Len:   106  Check: 9106  Weight:   1.00
Name: D_CD_84ZR0       Len:   106  Check: 5829  Weight:   1.00
Name: D_CD_ELI_K       Len:   106  Check: 2669  Weight:   1.00
Name: D_CD_NDK_M       Len:   106  Check: 4007  Weight:   1.00
Name: D_UG_94UG1       Len:   106  Check: 2098  Weight:   1.00
Name: F1_BE_VI85       Len:   106  Check: 4208  Weight:   1.00
Name: F1_BR_93BR       Len:   106  Check: 4525  Weight:   1.00
Name: F1_FI_FIN9       Len:   106  Check: 5556  Weight:   1.00
Name: F1_FR_MP41       Len:   106  Check: 4332  Weight:   1.00
Name: F2_CM_MP25       Len:   106  Check: 5691  Weight:   1.00
Name: F2KU_BE_VI       Len:   106  Check: 4047  Weight:   1.00
Name: G_BE_DRCBL       Len:   106  Check: 3839  Weight:   1.00
Name: G_NG_92NG0       Len:   106  Check: 825   Weight:   1.00
Name: G_SE_SE616       Len:   106  Check: 4456  Weight:   1.00
Name: H_BE_VI991       Len:   106  Check: 2728  Weight:   1.00
Name: H_BE_VI997       Len:   106  Check: 3468  Weight:   1.00
Name: H_CF_90CF0       Len:   106  Check: 5568  Weight:   1.00
Name: J_SE_SE702       Len:   106  Check: 4413  Weight:   1.00
Name: J_SE_SE788       Len:   106  Check: 3659  Weight:   1.00
Name: K_CD_EQTB1       Len:   106  Check: 4999  Weight:   1.00
Name: K_CM_MP535       Len:   106  Check: 4729  Weight:   1.00
Name: N_CM_YBF30       Len:   106  Check: 8961  Weight:   1.00
Name: O_CM_ANT70       Len:   106  Check: 7210  Weight:   1.00
Name: O_CM_MVP51       Len:   106  Check: 8490  Weight:   1.00
Name: O_SN_99SE_       Len:   106  Check: 726   Weight:   1.00
Name: O_SN_99SE_       Len:   106  Check: 9615  Weight:   1.00
Name: U_CD___83C       Len:   106  Check: 6063  Weight:   1.00

            1                                                  50
00BW0762_1  ........MI ELIAAVDYRI GVA.ALIIAL IIAIVVWTIA YIEYRKLLKQ
00BW0768_2  ........ML ELTARVDYRL GVG.ALIVAI ILAIVVWIWA YIEYKELLRQ
00BW0874_2  ........ML GLSEKAGYAL GVG.ALIVAL IIVIVVWTIV YIEYRKLVRQ
00BW1471_2  ........MI NLLERVDX.. GVG.ALGIAL IIVIVVWTIV YIEYRKLVRQ
00BW1616_2  .......... ..MLDLDYKV AVG.AFTVAL ILAIVVWILV YREYRKLLRQ
00BW1686_8  ...MEDVILS FIA.KIDYRI GIA.AIIVAL ILAIIVWTIV YLEYRKLVRQ
00BW1759_3  .......MID LSA.RVDYRI GVA.AFIIAL IIAIVVWTIV YIEYRKLLRQ
00BW1773_2  MLKLATIVDY ILAAKVDYRV GIG.ALIAAL IITIVVWIIV YREYRKLLRQ
00BW1783_5  ...IVDVIFS LTD.RVDYRI AVA.ALTIAL IIAIVVWTIV YIEYRKLVRQ
00BW1795_6  .......MVD WTKXKVDYRI AVV.AFIVAL IIAIVVWTIV YIEYRKLRKQ
00BW1811_3  .......... ..MLAIDYSL GVG.ALLVAL FIAIID.... ....RKLVRQ
```

```
00BW1859_5  ........ML ELTARVDYRL GVG.ALIVAL IIAIIVWTIA SLEYRKLKRQ
00BW1880_2  .......MLS LMT.RVDYRI AVA.AFVIAL ILAIIVWTIA YLEYRKLVKQ
00BW1921_1  .......MLD LAA.IVDYRI TIV.AFAIAL FIAIIVWTIA YLEYRKLVRQ
00BW2036_1  ........ML DLIAKVDYRV GIG.ALIVAL IIAVVVWIIA YIEYRKLLKQ
00BW2063_6  .......MID WTE.QVDYRI AIVXSFIVAL IIAIVIWTLA YIEYRKLSRQ
00BW2087_2  ........ML SLIERIDYRL GVG.ALIVAL IIVIIVWTIV YIEYRKLVRQ
00BW2127_2  ........ML DLAARVDYRL GVG.ALVVAL IIAIIVWTIV YIEYRKLVRQ
00BW2276_7  .........M VDLGRVDYRL GVG.ALIVAL IIAIVVWIIV YIEYRKLVRQ
00BW3819_3  ........ML DLLTRVDYRL GVG.ALIVAL IIAIIVWTIA YIEYRKLLRQ
00BW3842_8  ........MF DLLAGVDYRL GVG.ALIIAL IIAIVVWVIA YIEYKKWLKQ
00BW3871_3  .......MVD LLE.KVDYRI GIA.AFTVAL LIAIIVWIIA YIEYRKLVRQ
00BW3876_9  .........M LDLTQIGYEL GIG.ALIVAL IIAIVVWTIV YIEYRKVLRQ
00BW3886_8  .......... .MLAAVDYRL GVG.ALTVAL ILVIIVWTIV YIEYRKLVRQ
00BW3891_6  ........ML DLLAGVDYRI GVG.AFLVAL SIAIVVWTIV YIEYRKLLRQ
00BW3970_2  ........MF SLLERIDYRL GVG.ALLVAL IIAIVVWAIV YIEYRKLVRQ
00BW5031_1  .........M FALFEVDYRL TIG.AFIVAL FLAIVVWTIA YLEYRKLVRQ
 96BW01B21  ........ML ELIAKIDYRL GGG.ALIVAL SIAIVVWIIA YIEYKKLIRQ
  96BW0407  ........ML SLAA.IDYRI GVG.AFVVAL IIAIIVWIIV YIEYRKLVRQ
  96BW0502  ........MI NFLAKVDYRL GVG.ALIVAF IIAIVVWIIA YIEYRKLLRQ
 96BW06_J4  .......MID LLA.RVDYRI GLA.AFVVAL LIAIIVWTIV YLEYRKLVRQ
 96BW11_06  .......MVD LLA.KVDYKI AVA.AFIIAL IIAIVVWIIV YVEYRKLVKQ
  96BW1210  ........ML YLLEKVDYRL GVG.ALIIAL IIAIIVWTIA YLEYRKVLRQ
 96BW15B03  .......MVN FLAAKVDYRL GVG.ALIVAL IIAIVVWSIV YIEYKKLLKQ
 96BW16_26  ........MF SLLAKVDYRV GVG.ALIVAL IIAIIVWIIV YIEYRKLLRQ
 96BW17A09  ........ML NLLAKVDYRL GVG.ALVIAL IIAIVVWIIA YIEYRKLVRQ
 96BWMO1_5  .......MID WTA.RVDYRV AVV.AFIVAL ILAIIVWIIA YLEYRKLSKQ
 96BWMO3_2  .........ML SLAALVDYRL GVG.ALIVAL IIAIIVWTIV YLEYRKLVRQ
98BWMC12_2  ........MV VLGEKEIYIL GIG.ALIVAL IIAIVVWSIA YIEYRKLVRQ
98BWMC13_4  .......MSN WTA.RVDYRI AIA.AFIVAL IIAIVVWTIV YIEYRKLSRQ
98BWMC14_a  .......MLD LTE.GVDYRI GIA.AFAVAL FIAIIVWIIV YLEYRKIVKQ
98BWMO14_1  ........ML ELLEKVDYKI TVA.AFIVAL IIAIVVWTIA YIEYRKLLRQ
98BWMO18_d  ........ML DLIARVDYRL GVG.ALIVAL IIAIIVWTIV YIEYRKLVKQ
98BWMO36_a  ........ML AFLARVDYRL GVG.AFIIAL IIAIIVWTIA YLEYRKLVRQ
98BWMO37_d  ........ML NLVARVDYRL GVG.ALIVAL IIAIVVWTLV YIEYRKLVRQ
99BW3932_1  .......... .MLENIDYRL GVG.ALIVAL IIVIIVWTIV YIEYRKLVRQ
99BW4642_4  ........ML DLTARVDYRL GVG.ALIVAL IIAIVVWTIV YLEYRKLQRQ
99BW4745_8  ........MI NLIARVDYRL GVG.ALIIAL IIAIVVWTIA YIEYRKLLKQ
99BW4754_7  .......... .ML.ELDYRI GIA.AFIIAL IIAIVVWTIV YIEYRKLRRQ
99BWMC16_8  ........MF NLAARVDYRL GVG.ALVIAG IIAIVVWTIV YIEYRKLSRQ
A2_CD_97CD  .......... ..MSPLAILS IVG..LVVAS ILAIVVWTVV FIEYRKIKKQ
A2_CY_94CY  .......... ..MLPLVILA IVG..LIVAL ILAIVVWTIV FIEYKKIKKQ
A2D___97KR  .......... ..MNSLQILS IVG..LIVAF ILAVVVWTIV FIEYKKIKKQ
A2G_CD_97C  .......... ..MQALEISA IVG..LIVAF IAATVVWTIV YIEYRKIRKQ
A_BY_97BL0  .......... ..MTPLXIYA IVA..LVVVF VIXXVVWTIV GIXYXXLLKQ
A_KE_Q23_A  .......... ..MSPLEISA IVG..LIVAL ILAIAVWTIV GIEYKKIRKQ
A_SE_SE659  .......... ..MQSLVILA IVA..LVVVL IIAIVVWTIV FIEWGRLRRQ
A_SE_SE725  .......... .MLPLSTIWA IIG..LIVAL ILAIVVWTIV GIEYKKLLKQ
A_SE_SE753  .......... ..MIPLQIWA IAG..LIVAL ILAIIAWTIV GIEYEKLLKQ
A_SE_SE853  .......MSA LEISALEIWS IVG..LVVAL ILAIVVWTIV GIECKRLQKQ
A_SE_SE889  .......... ..MTPLEIWA IVG..LVVVS IIAIVVWTIV GIEYKKLLKQ
A_SE_UGSE8  .......... ..MSPLEIWA IVG..LIVAL ILAIVVWTIV GLEYKKLKRQ
A_UG_92UG0  .......... ..MQLLEICA VVG..LVVAL IIAIVVWTIV GIEYKKLLKQ
A_UG_U455_  .......... ..MTPLEIWA ITG..LIVAL ILAIVVWTIV GIE.KKLLKQ
AC_IN_2130  .......... ..MLNLNYEL AVG.ALIVAL IIAIVVWTIV YVEYRKLLRQ
AC_RW_92RW  .......... ..MTSLEIYA IVA..LIVAL IIVIVVWTLA GIEYKKLLKQ
AC_SE_SE94  ........ML DLLAEIDYRI GVG.AFIIAL IIAVVVWIIA YIEYRKLVKQ
ACD_SE_SE8  .......... ..MQPLEICS IVG..LIVAI ILAIVAWTIV GIEIKKLLRQ
ACG_BE_VI1  ........ML ELLAQVDYRI GVG.ALILAL FIAIVVWIIA YLEYRKLVRQ
AD_SE_SE69  .......... ..MQSLVILS IVA..LVVAL VIAIVVWTIV FIEYRRLERQ
AD_SE_SE71  .......... .MLSLLHICA IVG..LLVAL IIAIVVWTIV GIEYKRLLRQ
ADHK_NO_97  .......... ....MYILGI GIG.ALIVAF IIVIVVWTIV YLEYRKLVKQ
```

```
ADK_CD_MAL  .......... ..IQPLVILA IVA..LVVTL IIAIVVWTIV FIEIRKIRRQ
AG_BE_VI11  .......... ..MQSLEIAA VVG..LIVAF IAAIIVWTIV FIEYRKIRKQ
AG_NG_92NG  .......... ..MQSLEIAA IAG..LVVAA IAAIVVWTI. .....KIKKQ
AGHU_GA_VI  .......... ....MYILGI GIG.ALVVAF IIAIVVWTIA YIEYRKLLRQ
AGU_CD_Z32  .......... ..MQFLEIAS IVG..LIVAV LAAIVVW... .IEYRKIRKE
AJ_BW_BW21  .......... ..MRPLQIAA IVG..VIVAI FLAIVVWTIV YIEYKRLVRQ
B_AU_VH_AF  .......... ..MQSLQIVA IVA..LVVVT IIAIVVWTIV FIEYRRILRQ
B_CN_RL42_  .......... ..MQALTILA IVA..LVVVA IIAIVVWTIV LLEYRKILRQ
B_DE_D31_U  .......... ..MQSLQVAA IVA..LVVVA IIAIVVWTIV FIEYRKILRQ
B_DE_HAN_U  .......... ..MLSLQIAA IVA..LVVAA IIVIVVWTIV VIEYRKLLRQ
B_FR_HXB2_  .......... ..TQPIPIVA IVA..LVVAI IIAIVVWSIV IIEYRKILRQ
B_GA_OYI_M  .......... ..IYSLQILA IVA..LVVVT IIAIVVWTIV LLEYRKILRQ
B_GB_CAM1_  .......... ...MLPLQIA IVA..LVVVA IIAIVVWTIV FIEYRKIRRQ
B_GB_GB8_A  .......... .....IQILT IVA..LVVAG IVAIVVWIIV FIEYRKILKK
B_GB_MANC_  .......... ..IVSLQIVA IVA..LVVVT IIAIVVWTIV FIEYRKILRQ
B_KR_WK_AF  .......... ..MQLLAILA IVG..LVVAA ILAIVVWFIV FIEYKKILKQ
B_NL_3202A  .......... ..MQPLVVAA IVA..LVVVA IIAIVVWSIV FIEYRKILRQ
B_TW_TWCYS  .......... ..MQPLHILA IVA..LVVAA IIAIVVWTIV LLEYRKILRQ
B_US_BC_L0  .......... ..MQPIQIAA IVA..LVVVG IIAIVVW... .....KILRQ
B_US_DH123  .......... ..MQPLVILA IVA..LVVAL IIVIVVWSIV LIEYRKILRQ
B_US_JRCSF  .......... ..MQPLQILA IVA..LVVAG IIAIIVWSIV LIEYRKILRQ
B_US_MNCG_  .......... ..MQPLVIAA IVA..LVVAG IIAIVVS.IV FIEYRKIRRQ
B_US_P896_  .......... ..MLSLQILA IVA..LVVAA IIAIVVWSIV FIEYRKILRQ
B_US_RF_M1  .......... ..MQSLEILA IVA..LVVAA ILAIVVWTIV GIEIRKTLRQ
B_US_SF2_K  .......... ..MQSLQILA IVS..LVVVA IIAIVVWTIV LIEYRKILRQ
B_US_WEAU1  .......... ..MQSLQILA IVA..LVVAG IIAIVVWSIV LIEYRKILRQ
B_US_WR27_  .......... ...MPLYILA VVA..LVLAA IIAIVVWTIV FIEYRKILRQ
B_US_YU2_M  .......... ..LQSLQVLA IVA..LVVAT IIAIVVWTIV FIEYRKILRQ
BF1_BR_93B  .......... ....MSYLLV IGLAALIAAL IIAIVVWTIA YIEYRELVRQ
C_BR_92BR0  ........ML ELIGRIDYRL GVG.ALIVAL IIVIIVWTIA YIEYRKLVRQ
C_BW_96BW0  ........MI SLAARVDYRI GVG.AFIVAL IIAIIVWTIV YIEYRKLVRQ
C_BW_96BW1  .......MVD LLA.KVDYKI AVA.AFIIAL IIAIVVWIIV YVEYRKLVKQ
C_BW_96BW1  ........ML YLLEKVDYRL GVG.ALIIAL IIAIIVWTIA YLEYRKVLRQ
C_BW_96BW1  .......MVN FLAAKVDYRL GVG.ALIVAL IIAIVVWSIV YIEYKKLLKQ
C_ET_ETH22  ........MV DLLAKVDYRI VIV.AFIVAL IIAIVVWTIA YIEYRKLLRQ
C_IN_93IN1  .......... ...MIDLYRL GVG.ALIVAL IIAIVVWTIV YIEYRKLVRQ
C_IN_93IN9  .......... ..MLDLDYKL AVG.ALIVAL IIAIVVWIIA YIEYRKLVKQ
C_IN_93IN9  .......... ..IIDLDYRL GVG.ALIVAF IIAIVVWTIV YIKYRKLLRQ
C_IN_94IN1  .......... ....MLDLKL AVG.ALIVAL IIAIVVWTIV YIEYRRLVKQ
C_IN_95IN2  .......... ..MVNLDYKL GVG.ALIVAL IIAIVVWTIV YIEYRKLVQQ
CRF01_AE_C  .......... ...MSALQIA IVG..LIVAL ILAIVVWTIV FIEYKKILRQ
CRF01_AE_C  .......... ..RLPLHICA IVG..LIVAL IIAIVVWTIV AIEYKKLRRQ
CRF01_AE_C  .......... ..MLPLQISA IVR..LIVAL ILAIVVWTLV AIEYRILR.Q
CRF01_AE_T  .......... ..MSPLEISA IVG..LIVAI ILVIVVWTIV AIEFKKILGQ
CRF01_AE_T  .......... ..MSPLEISA IVG..LIVAL ILAIVVWTIV AIEFKKILRQ
CRF01_AE_T  .......... ...MHLWKLA IVG..LLVAL ILAIVVWTIV AIEFKKILRQ
CRF01_AE_T  .......... ..MTPLEISA VVG..LIVVL ILAIVVWTIV GILKKILR.Q
CRF01_AE_T  .......... ..MTPLEISA IVG..LIVAL ILAIVVT.IV AIEVKKILRQ
CRF01_AE_T  .......... ..MTPLQISA IVG..LIVAL ILAIVVT.IV IIEYKKILKQ
CRF02_AG_F  .......... ..MQSLEITA IVG..LVVAF IAAIVVWAIV FIEYRKIRKQ
CRF02_AG_F  .......... ..MKSLEIAA IVG..LVVAI IAGIVVWTIV VIEYRKIRKQ
CRF02_AG_G  .......... ..MQSLEIAA IVG..LVVAF IAAIVVWIIV YIEYRKIRKQ
CRF02_AG_N  .......... ..MQPLTITA IVG..LVVAF IAAIVVWSIV YIEYRKIRKQ
CRF02_AG_S  .......... ..MNSLDIVA IVG..LVVAF IAAIVVWTIV YIEYRKIRKQ
CRF02_AG_S  .......... ..MQPLAVAA IVG..LVVAF ILAIVVWTIV FIEYRKIRKQ
CRF03_AB_R  .......... ..MQSLAIAA IVA..LVVVG IIAIVVGSIV FIEYRKILRQ
CRF03_AB_R  .......... ..MQFLVIAA IVA..LVVGG IIAIVVWSIV FIEYRKILRQ
CRF04_cpx_  .......... ..MLFWEIWA IVG..LVVAL IIVIVVWTLV FIEYKKLRRQ
CRF04_cpx_  .........M SSLSSLEIWT IVA..LIVAL IIAIVVWTIV YIEYKKLQRQ
CRF04_cpx_  .........M SSMSSLQVWA ITA..LIVAL IIAIIVWTIA LIEYRKLLRQ
CRF05_DF_B  .......... ....MSDLLT IAVVALIVAL IIAIVVWIIV YIEYKKLVRQ
```

```
CRF05_DF_B  ..........  ....MSDLLA VAIAAFIVAL IIAIVVWTIV YLEYRKLVRQ
CRF06_cpx_  ..........  ..MQALEIAA IVG..LVVAF LAAIVVWTIV FIQYREIRKQ
CRF06_cpx_  ..........  ..MQSLEISA IVG..LVVAF LAAIVVWTIV FIQYREIRKQ
CRF06_cpx_  ..........  ..MQPLEISA VVG..LIVAF LAAIVVWTIV YIQYREIRKQ
CRF06_cpx_  ..........  ..MQPLEISA IVG..LIVAF LAAIVVWTIV FILHRQIKKQ
CRF11_cpx_  ..........  ..MIPLQIAG IVA..LIVAA IIVIIVWTIV FIEYRKLLKQ
CRF11_cpx_  ..........  ..MITLRTVK HSS..ISQHL LIAIIVWTIA GIEYRRLLRH
D_CD_84ZR0  ..........  ..MQSLQILA IVA..LVLAL IIAIVVWTIV FIEYRRIKRQ
D_CD_ELI_K  ..........  ..MQPLGIIA IAA..LVVAI ILAIVVWTIV FIEYRRIKKQ
D_CD_NDK_M  ..........  ..MQPLVIIA IAA..LVVAI IIAIVVWTIV YIEYRRIKRQ
D_UG_94UG1  ..........  ..MQPLEILA IVA..LVVAL ILAIVVWTIV FIECKKLRRQ
F1_BE_VI85  ..........  ....MSYLLA IGIAALIVAL IIAIVVWTIV YIEYKKLVRQ
F1_BR_93BR  ..........  ....MSNLLA IGIAALIVAL IITIVVWTIA YIEYKKLVRQ
F1_FI_FIN9  ..........  ....MSDLLA ITIVAFIVAL IIVIVVWTIV FIEYKKLVRQ
F1_FR_MP41  ..........  ....MSNLYV LSIVAFIIAL IIAIVVWTIV FIEYKKLLRQ
F2_CM_MP25  ..........  ....MSLSLI VVIAAYIVVL ILAIIVWTIV YIEYKKILRQ
F2KU_BE_VI  ..........  ....MNL.LL VGIGALIVAF LLAIVVWTIA YLEYRKVLKQ
G_BE_DRCBL  ..........  ..MQPLEISA IVG..LIVAS IAAIVVWTIV FIEYRKIRKQ
G_NG_92NG0  ..........  ..MQALEIS. .....LIVAF IAATIVWSIV FIEYRKIRKQ
G_SE_SE616  ..........  ..MQSLVILA IVG..LIVAF IAAIVVWTIV FIEYREIRKQ
H_BE_VI991  ..........  ....MNILGI GIG.ALVVAF IIAIVVWTIA YIEYRKLK.Q
H_BE_VI997  ..........  ....MYIIGI GIG.ALIVAF IIAIVVWTIV YIEYRKLVKQ
H_CF_90CF0  ..........  ....MYILGL GIG.ALVVTF IIAVIVWTIV YIEYKKLVRQ
J_SE_SE702  ..........  ..MVSLQIVA IVA..LIVAF FLAICVWTIV YIEYKKLLRQ
J_SE_SE788  ..........  ..MIPLQIAA IVA..FIVAI FLAIGMWTIV YIEYKKLLRQ
K_CD_EQTB1  ..........  .....MVPLT VGIIALVAAL ILAIIVWTIA YLEYRKVVRQ
K_CM_MP535  ..........  .....MVSLA ISIVALVVAL ILAIIVWTIV YIEYRKLVKQ
N_CM_YBF30  ..........  .......MLS LGFIALGAAV SIAVIVWALL YREYKKIKLQ
O_CM_ANT70  ..........  ..MHHRDLLA III..ISALL FINVILWGFI LRKYLEQKEQ
O_CM_MVP51  ..........  ..MHQENLLA LIA..LSALC LINVLIWLFN LRIYLVQRKQ
O_SN_99SE_  ..........  ..MHHRDLLA LIT..TSALL LTNVVLWTFI LRQYLKQKKQ
O_SN_99SE_  ..........  ..MHHRDLLT LII..ISALL LTNVILWAFI LRQYLRQKKQ
U_CD___83C  ..........  ..MSILQIVA IVA..IIVAL ILAIVVWTIV YIEYKRLLRQ

            51                                                  100
00BW0762_1  KKIDR....L IKRIIERAED SGNESDGD.H EEL....STM VDMGHLRLLD
00BW0768_2  RKIDW....L IKRIRERAED SGNESEGD.T EEL....STM VDMGQLRLLD
00BW0874_2  RKID.....L IKRIRERAED SGNESDGD.T EEL....STM VDMGNLRLLD
00BW1471_2  KKIDQ....L VKRIREREED SGNESDGD.I EEL....STM VDMGHLRLLD
00BW1616_2  KRIDN....L IKRIRERAED SGNESDGD.T EEL....STL VDMGNLRLLD
00BW1686_8  RKIDR....L IERIKEREED SGNESEGD.T EEL....ATM VDMGQLRLLD
00BW1759_3  RKINW....L IERIRERAED SGNESEGD.T EEL....ATM VDMGHLRLLH
00BW1773_2  KKIDR....V LEKIRERAED SGNESEGD.L EEL....STL VDVEHLRLVD
00BW1783_5  RKIDR....L IERIRERAED SGNESDGD.Q EEL....STM VDMGHLRLLD
00BW1795_6  RKINW....L IERIRERAED SGNESDGD.Q EEL....ATM VDMGHLRLLD
00BW1811_3  AKIDK....I IKRIRERAED SGNESEGD.I EEL....STM VDMGHLRLMD
00BW1859_5  RKIDW....L IKRIRERAED SGNESEGD.T EEL....ATM VDMGHLRLLD
00BW1880_2  KKIDW....L IERIRERAED SGNESDGD.H EEL....STM VDMGHLRLLD
00BW1921_1  RKIDR....L IKRIRERAED SGNESEGD.T EEL....ATM VDMGQLRLLG
00BW2036_1  RKIDW....L IKRIRERAED SGNESEGD.I EEL....ATM VDMGHLRLLD
00BW2063_6  RKIDW....L IKRIRERAED SGNESDGD.H EEL....ATM VDLGHLRLLD
00BW2087_2  RKLDW....L VKRIRERAED SGNESDGD.V EEL....STM VDMGHLRLLD
00BW2127_2  RKIDW....L IKRIGERAED SGNESEGD.T EEL....ATM VDMGHLRLLD
00BW2276_7  RKIDW....L VKRIRERAED SGNESEGD.T EEL....ATM VDMGHIRLLD
00BW3819_3  RKIDW....L IKRIREREED SGNESEGD.T EEL....AAM VDMGHLELLG
00BW3842_8  KKIDW....L IERIRERAED NGNESEGD.T EEL....ATM VDMGHLRLLD
00BW3871_3  RKIDW....L IKRIRERAED SGNESEGD.T EEL....ATM VDMGQLRLLD
00BW3876_9  RKIDR....L IKRIRERAED SGNESEGD.I EEL....EAI VDMGHLRLLG
00BW3886_8  RKIDW....L IKRIRERAED SGNESEGD.I EDL....AMV VDMGQLRLLD
00BW3891_6  KRIDR....L IERIRERAED SGNESEGD.I EEL....STM VDMGHLRLLD
00BW3970_2  RRLDW....L VKRIRERAED SGNESEGD.N EEL....STV VDLGHLRLLD
```

```
00BW5031_1  RRIDQ....L VKRIRERAED SGNESDGD.T EEL....STM VDMGHLRLLD
 96BW01B21  RRINW....L IERIRERAED SGNESEGD.T EEL....STM VDMGHLRLLD
  96BW0407  RKIDW....L VKRIRERAED SGNESDGD.T EEL....STM VDMGHLRLLD
  96BW0502  RKIDR....L IRKIRERAED SGNESDGD.I EEL....SAM VDVGHLRLLD
 96BW06_J4  RKIDW....L IKRIRERAED SGNESEGD.T EEL....ATM VDMGQLRLLD
 96BW11_06  RKIDW....L IERIRERAED SGNESDGD.H EEL....STM VDMGHLRLLD
  96BW1210  RRIDR....L VERIREREED SGNESEGD.I EEL....STM VDMGHLRLLD
 96BW15B03  KNIDW....L IKRIRERAED SGNESEGD.T EEL....ATM VDMGHLRLLD
 96BW16_26  KKIDR....L IERIRERAED SGNDSEGD.I EDL....STM VDMDHLRLLD
 96BW17A09  RKLDE....L VKRIRERAED SGNESDGD.I EEL....STM VDMGHLRLLD
 96BWMO1_5  RKIDC....L IKRIRERAED SGNESDGD.Q EEL....ATM VDMGRLRLLD
 96BWMO3_2  RKINW....L VKRIRERAED SGNESEGD.T EEL....ATM VDLEHFGLLD
98BWMC12_2  RRIDQ....L IKRIGERAED SGNESDGD.T EEL....STL VDMGHLRLLD
98BWMC13_4  RKIDR....L IKRIRERAED SGNESDGD.Q DEL....ATM VDMGHLRLLD
98BWMC14_a  RKIDW....L IERIRERAED SGNESEGD.T EEL....ATM VDMGQLRLLD
98BWMO14_1  RKIDC....L IKRIRERAED SGNESDGD.T EEL....STM VDMGNLRLLD
98BWMO18_d  RKIDW....L IKRIRERVED SGNESDGD.T EEL....STL VDMGNLRLLD
98BWMO36_a  KKIDR....L IERIRERAED SGNESEGD.T EEL....STM VDMGHLRLLD
98BWMO37_d  RKIDQ....L VKRIRERAED SGNESEGD.I EEL....STM VDMGNLRLLD
99BW3932_1  RKIDW....L IKRIRERAED SGNESEGD.T EEL....ATM VDMEHLRLL.
99BW4642_4  RRIDW....L IKRIRERAED SGNESEGD.T EEL....ATM VDMGHLRLLD
99BW4745_8  KKIDW....L IKRIRERAED SGNESEGD.T EGL....STM VDMEHLRLLD
99BW4754_7  RKIDR....L IERIRERAED SGNESDGD.Q EEL....SAM VDMGHLRLLD
99BWMC16_8  RKINE....I IKRIRERAED SGNESDGD.M EEL....STM VDLGHLGLLD
A2_CD_97CD  RKIDW....L LERISERAED SGNESDGD.T EEL....SKM VGMGNLGFWD
A2_CY_94CY  RKIDW....L IKRISERAED SGNESDGD.T EEL....SAL VERGHLDFGD
A2D___97KR  NKIDW....L IKRISERAED SGNESEGD.T EEL....STL MEMGNLDFGD
A2G_CD_97C  KRIER....L LDRIGERAED SGNESEGD.A EEL....AKL VEMGGFDPWV
A_BY_97BL0  RKIDR....L XERXXXRAED SGNESEGD.A XEL....STL MEVXNYALLD
A_KE_Q23_A  RKIDR....L IERIRERAED SGNESDGD.T EEL....STL IDMGDYDLGD
A_SE_SE659  R.KDR....L IERIRERAED SGNESDGD.T EEL....SAL IEMGNYNLGF
A_SE_SE725  RKIDK....L VERIRERAED SGNESDGD.T NEL....AAL IEMGNYNLGD
A_SE_SE753  KEIDG....L VERIAERAED SGNESDGD.T EEL....STL IEMGNYDLGV
A_SE_SE853  RKIDR....L IERIRERAED SGNESEGD.T EEL....SAL IEMGNYDLGN
A_SE_SE889  RKIDR....L IRRISERAED SGNESDGD.T DEL....AKL VEMGNYDLGD
A_SE_UGSE8  RKIDR....L IERIRERAED SGNESDGD.T EEL....STL IELGDYDLGD
A_UG_92UG0  RKIDR....L VDRIRERAED SGNESDGD.R EEL....SLL VDMGDYDLGD
A_UG_U455_  KKIDR....L LNRIRERAED SGNESDGD.T EEL....SLL VEMGNYDLGV
AC_IN_2130  RKIDR....L IKRIRERAED SGNESEGD.T EEL....STM VDMGQLRLLD
AC_RW_92RW  RKIDR....L IKKIRERAED SGNESDGD.I DEL....SKL VGVGNYDLGD
AC_SE_SE94  RKIDW....L IKRIRERAED SGNESDGD.T EEL....STM VDMGHLRLLD
ACD_SE_SE8  KKIDR....L IERIRERAED SGNESDGD.T EEL....AAL VEMGNYDPGD
ACG_BE_VI1  RKIDW....L VKRIRERAED SGNESEGD.T EEL....STM VDMGELRLMD
AD_SE_SE69  KRIDW....L IDRIRERAED SGNESEGDRD ELS....EL. VEMGHPAPWD
AD_SE_SE71  RKIDR....I LDRIRVRAED SGNESDGD.T EEL....SKL IDMGNYDLGD
ADHK_NO_97  KRIDQ....L IERIRERAED SGNESDGD.T EEL....SEL VEMGRLNLGY
ADK_CD_MAL  RKIDR....L IDRIRERAED SGNESEGDTE ELS....KL. VEMGHDAPWD
AG_BE_VI11  KKIEK....L PDRIRERAED SGNESEGD.T DEL....ATL VERGNFDPWV
AG_NG_92NG  EKIDR....L LDRIRERAED SGNESEGD.T EEL....ATL VDMVDFDPWV
AGHU_GA_VI  RKIDW....L IRRISERAED SGNESDGD.T EEL....STL VEMGNHNPGY
AGU_CD_Z32  KKIEY....L IDRIRKRAED SGNESEGD.T EEL....STL MEMGDFHPWV
AJ_BW_BW21  KKIDR....I IDRIGERAED SGNESDGD.T DEL....HKL VEMGPHDLWN
B_AU_VH_AF  RKIDR....L IDRIAERAED SGNESEGD.Q EEL....AAL VEMGHHAPWD
B_CN_RL42_  RKIDR....L IDRIRERAED SGNESDGD.Q EEL....SAF MEMGHHAPWD
B_DE_D31_U  RKIDR....L IDRIIERAED SGNESEGD.Q EEL....SAL VEMGHHAPWD
B_DE_HAN_U  RKINR....L IDRLTERAED SGNESEGD.Q EEL....SAL VERGHLAPWD
B_FR_HXB2_  RKIDR....L IDRLIERAED SGNESEGEIS ALV....EMG VEMGHHAPWD
B_GA_OYI_M  RKIDR....L IDRIRERAED SGNDSEGD.Q EEL....SAL VEMGHHAPWD
B_GB_CAM1_  KQVDR....L IDRIIERAED SGNESEGD.Q EEL....SAL MEMGHNAPWD
B_GB_GB8_A  KKIDR....L IDRIIERAED SDNESEGD.Q EEL....SAL VEMGHHAPWD
B_GB_MANC_  RKIDR....L IDRIRERAED SGNESEGD.Q EEL....SEL VEMGHLVPWD
```

```
B_KR_WK_AF   KKIDR....L IDRIRERAED SGNESEGD.Q EEL....SAL VEMGHHAPWD
B_NL_3202A   RKIDR....L IDRIAERAED SGNESEGD.Q EEL....SAL VEMGHHAPWD
B_TW_TWCYS   RKIDR....L IDRIRERAED SGNESEGD.Q DEL....SAL VEMGYHAPWD
B_US_BC_L0   RKIDR....L IKRISKRAED SGNESEGD.Q EEL....SAL VEMGHHALWD
B_US_DH123   KKIDR....L IDRIRERAED SGNESDGD.Q EEL....SAL VERGHLAPWD
B_US_JRCSF   RKIDR....L IDKIRERAED SGNESEGD.Q EEL....SAL VERGHLAPWD
B_US_MNCG_   RKIDR....L IDRISERAED SGNESEGD.Q EEL....SAL VGMGHDAPWV
B_US_P896_   RKIDR....L IDRIREREED SGNESEGD.Q EEL....AAL ERG.HLAPWD
B_US_RF_M1   KKIDR....L IDRIRERAED SGNESDGD.E EEL....SAL VEMGHHAPWD
B_US_SF2_K   R.IDR....L IDRIREKAED SGNESEGD.Q EEL....SAL VEMGHLAPWD
B_US_WEAU1   RKIDR....L IDRIRDRAED SGNESEGD.Q EEL....SAL VEMGHHAPWD
B_US_WR27_   RKIDR....L IDRIRERAED SGNESEGD.Q EEL....SAL MEMGHHAPWD
B_US_YU2_M   RKIDR....L INRITERAED SGNESDGD.Q EEL....SAL VERGHLAPWD
BF1_BR_93B   RKINR....L YKRIRERAED SGNESEGD.A EEL....AAL GEMGPFIPGD
C_BR_92BR0   RRIDW....L VKRIKERAED SGNESGGD.T EEL....ETM VDMGHLRLLD
C_BW_96BW0   RKIDW....L VKRIRERAED SGNESDGD.T EEL....STM VDMGHLRLLD
C_BW_96BW1   RKIDW....L IERIRERAED SGNESDGD.H EEL....STM VDMGHLRLLD
C_BW_96BW1   RRIDR....L VERIREREED SGNESEGD.I EEL....STM VDMGHLRLLD
C_BW_96BW1   KNIDW....L IKRIRERAED SGNESEGD.T EEL....ATM VDMGHLRLLD
C_ET_ETH22   RRIDR....L IKRTRERAED SGNESDGD.T EEL....STM VDMGNLRLLD
C_IN_93IN1   SKINW....L IKRIRERAED SGNESEGD.T EEL....STM VDMGRLRLLD
C_IN_93IN9   RKIDW....L IKRIRERAED SGNESEGD.T EEL....STM VDMGRLRLLD
C_IN_93IN9   RRIDW....L IKRIRERAED SGNESEGD.T EEL....STM VDMGHLRLLD
C_IN_94IN1   RKIDW....L IKRIRERAED SGNESEGD.T EEL....STM VDMGRLRLLD
C_IN_95IN2   RKIDW....L IKRIRERAED SGNESEGD.T EEL....STM VDMGRLRLLD
CRF01_AE_C   RKIDR....L IERIRERAED SGNGSEGD.T NEL....ATL VEVGDFDPWV
CRF01_AE_C   RKIDR....L VQRISERAED SGNESEGD.T EEL....AKL VEMGDFDPWV
CRF01_AE_C   RKIDR....L IERIRERAED SGNESEGD.T DEL....AKL VEMGDFDPWV
CRF01_AE_T   RKIDR....L VKRIRERAED SGNESEGD.T DEL....AKL VEMGDFDPWV
CRF01_AE_T   RKIDR....L VKRIRERAED SGNESEGD.T DEL....AKL VEMGDFDPWV
CRF01_AE_T   RKIDR....L VKRIREREED SGNESEGD.T DKL....AKL VEMGDFDPWV
CRF01_AE_T   RKIDR....L VKRIRERAED SGNESEGD.T DEL....AQL VEMEDFDPWV
CRF01_AE_T   RKIDR....L VKRIRERAED SGNESEGD.T DEL....AKL VEMGDFDPWV
CRF02_AG_F   KKIDK....L IKRIGERAED SGNESEGD.T DEL....AKL VEMGDFDPWV
CRF02_AG_F   KKIDK....L LDRIRERAED SGNESDGD.A EEL....STL MEMGYD.HIL
CRF02_AG_F   KKIDK....L LDRIRERAED SGNESDGD.T EEL....STL LEMGYD.NIL
CRF02_AG_G   KKIDK....L LDRIREREED SGNESEGD.A EEL....SKL MEMGHD.FWI
CRF02_AG_N   KKIDR....L LDRIRERAED SGNESDGD.T EEL....STL MEMGYE.YIL
CRF02_AG_S   KKIDR....L LDRIRERAED SGNESDGD.T EEL....STL MEMGYD.NIL
CRF02_AG_S   GKIDK....L LDRIRERAED SGNESDGD.T EEL....STL LEMGYDNAAL
CRF03_AB_R   RKIDR....L IDRIRERAED SGNESEGD.Q E.......AL MEMGHLVPWD
CRF03_AB_R   RKIDR....L IDRIRERAED SGNESEGD.Q E.......AL MEMGHLAPWD
CRF04_cpx_   RRIDS....L YNRIRERAED SGNESDGD.A EEL....STL VGMGNFDPWV
CRF04_cpx_   RKIDR....L YKRIRERAED SGNESDGD.T EEL....STL VGMGDFDPWV
CRF04_cpx_   RKIDR....L CKRIIERAED SGNDSDGD.T EEL....STL VDMGDFHPLV
CRF05_DF_B   RKINR....L YKRIRERAED SGNESEGD.A EEL....AAL GEVGPFIPGD
CRF05_DF_B   RKINR....L YKRIRERAED SGNESEGD.A EEL....AAL GEMGPFIPGN
CRF06_cpx_   KKIEK....L LDRIRERAED SGNESEGD.T DEL....ATL MEMGDFDPWV
CRF06_cpx_   RKIEK....L LNRIRERAED SGNESEGD.T EEL....AAF MEMGNFDPWV
CRF06_cpx_   KKIEK....L LDRIRERAED SGNESEGD.T DEL....ATL MEMGNFDPWV
CRF06_cpx_   KKIEK....L LDRIREREED SGNDSEGD.T EEL....ATL MEMGNFDPWV
CRF11_cpx_   KKIDR....L IDRIRERAED SGNESEGD.T EEL....ARL VEMGPHDQWN
CRF11_cpx_   R........K DRLRIRRAED SGNESEGD.T EEL....AQL VEMGPHDLWN
D_CD_84ZR0   RKIDW....L IDRIREREED SGNESEGDKE ELS....TL. VEMGHHAPWD
D_CD_ELI_K   RRIDC....L LDRITERAED SGNESEGDRE KLS....KL. VEMGHHAPWD
D_CD_NDK_M   RKIDC....L IDRIRERAED SGNESEGERE ELS....KL. VEMGHHAPWD
D_UG_94UG1   RKIDW....L IDRIRERAED SGNESEGDKE ELS....AL. VEMGHDAPWD
F1_BE_VI85   RKINK....L YKRIRERAED SGNESEGD.A EEL....AAL GEMGPFIPGD
F1_BR_93BR   RKINR....L YKRISERAED SGNESEGD.A EEL....AAL GEVGPFIPGD
F1_FI_FIN9   RKINR....L YIRIRERAED SGNESEGD.A EEL....AAL GKMGPFIPGD
F1_FR_MP41   RKINR....L YERIRERAED SGNESEGD.A EEL....AAL GEMGSFISGD
```

```
F2_CM_MP25   KRINR....L YERIIERAED SGNESEGD.A EEL....AAL GEVGPLIPGD
F2KU_BE_VI   ERINQ....L YNRLIERAED SGNESEGE.A EEL....AAL GEVGHLVLGN
G_BE_DRCBL   KRIEK....L LDRIRERAED SGNESEGD.T EEL....ATL MELGDFDPWV
G_NG_92NG0   KKIEK....L LDRIRERAED SGNESEGD.T EEL....ATL MEMGDFDPWV
G_SE_SE616   KRIGK....L LDRIRERAED SGNESDGD.T EEL....VTL VEMGDFDPWV
H_BE_VI991   RKIDR....L IERIRERAED SGNESDGD.T EEL....SKL VEMGHLNLGY
H_BE_VI997   KKIDR....L IQRIIEGAED SGNESD.... EEL....STM VERGHLTFGY
H_CF_90CF0   KKIDR....L IERIGERAED SGNESDGD.T EEL....SKL MEMGHLNLGY
J_SE_SE702   RKIDK....L INRIRERAED SGNESDGD.T DEL....AEL VEMGPHDLWN
J_SE_SE788   RKIDK....L IDRIRERAED SGNESDGD.T EEL....ADL VERGPHDLWN
K_CD_EQTB1   KRINW....L FDRIRERAED SGNESEGD.T EEL....AAL GETGHLILGD
K_CM_MP535   KRINW....L IDRIRERAED SGNESEGD.A EEL....ADI GELGHLILGN
N_CM_YBF30   EKIKH....I RQRIREREED SGNESDGD.A EWLDGDEEWL VTLLSSSKLD
O_CM_ANT70   DRKEREILER LRRIREIRDD SDYESNGE.. EEQ.....EV MDLVLSHGFD
O_CM_MVP51   DRREQEILER LRRIKEIRDD SDYESNEE.. EQQ.....EV MELIHSHGFA
O_SN_99SE_   DKREREILER LRRIRQIEDD SDYESDGT.. EEQ.....EV RDLVHSYGFD
O_SN_99SE_   DRREREILER LRRIRQIEDD SDYESDGK.. EEQ.....EV RDLVHGYGFD
U_CD___83C   RKIDW....L IDRIRERAED SGNESEGD.T EEL....STL VEMEPDNFRN

             101
00BW0762_1   ANGL..
00BW0768_2   GNDL..
00BW0874_2   VNDL..
00BW1471_2   VNDL..
00BW1616_2   DL....
00BW1686_8   VNVL..
00BW1759_3   DNNL..
00BW1773_2   INH...
00BW1783_5   AHDL..
00BW1795_6   ANNL..
00BW1811_3   IINY..
00BW1859_5   INDL..
00BW1880_2   ANDL..
00BW1921_1   HGL...
00BW2036_1   VHDL..
00BW2063_6   ANDL..
00BW2087_2   VNDL..
00BW2127_2   DL....
00BW2276_7   GNDL..
00BW3819_3   AHDL..
00BW3842_8   L.....
00BW3871_3   VNDI..
00BW3876_9   INNL..
00BW3886_8   VNNL..
00BW3891_6   VNDV..
00BW3970_2   VTDL..
00BW5031_1   VNDL..
 96BW01B21   DNAL..
  96BW0407   DI....
  96BW0502   VNN...
 96BW06_J4   NL....
 96BW11_06   ANDL..
  96BW1210   ADGL..
 96BW15B03   L.....
 96BW16_26   INN...
 96BW17A09   VNDL..
 96BWMO1_5   TNDL..
 96BWMO3_2   INL...
98BWMC12_2   DNEL..
98BWMC13_4   VNDL..
98BWMC14_a   VM....
```

98BWMO14_1 ANDL..
98BWMO18_d ANDL..
98BWMO36_a AHDL..
98BWMO37_d ANDL..
99BW3932_1 ......
99BW4642_4 VNDL..
99BW4745_8 DL....
99BW4754_7 VNDL..
99BWMC16_8 ANDL..
A2_CD_97CD DNDV..
A2_CY_94CY VNNV..
A2D___97KR AND...
A2G_CD_97C GDNL..
A_BY_97BL0 DNNV..
A_KE_Q23_A NNIL..
A_SE_SE659 DNNL..
A_SE_SE725 DNDL..
A_SE_SE753 GNNL..
A_SE_SE853 DNNL..
A_SE_SE889 NNNL..
A_SE_UGSE8 DNNL..
A_UG_92UG0 DNNL..
A_UG_U455_ DNNL..
AC_IN_2130 VNGL..
AC_RW_92RW VNNL..
AC_SE_SE94 VNNL..
ACD_SE_SE8 DINL..
ACG_BE_VI1 AIDL..
AD_SE_SE69 VDDM..
AD_SE_SE71 DNNL..
ADHK_NO_97 VADL..
ADK_CD_MAL VDDL..
AG_BE_VI11 GDNL..
AG_NG_92NG GDNL..
AGHU_GA_VI VNDL..
AGU_CD_Z32 GDNL..
AJ_BW_BW21 VNDL..
B_AU_VH_AF VDDL..
B_CN_RL42_ VDDL..
B_DE_D31_U VDDL..
B_DE_HAN_U VNDQ..
B_FR_HXB2_ VDDL..
B_GA_OYI_M VDDM..
B_GB_CAM1_ VNDL..
B_GB_GB8_A VDDL..
B_GB_MANC_ VDDL..
B_KR_WK_AF VDDL..
B_NL_3202A VDDL..
B_TW_TWCYS VNDQ..
B_US_BC_L0 IDDL..
B_US_DH123 IDDL..
B_US_JRCSF INDL..
B_US_MNCG_ INDL..
B_US_P896_ VDDL..
B_US_RF_M1 VDDL..
B_US_SF2_K VDDL..
B_US_WEAU1 IDDL..
B_US_WR27_ VDDL..
B_US_YU2_M VDDL..
BF1_BR_93B IDNL..
C_BR_92BR0 GNDL..

```
C_BW_96BW0   DN....
C_BW_96BW1   ANDL..
C_BW_96BW1   ADGL..
C_BW_96BW1   L.....
C_ET_ETH22   VNDL..
C_IN_93IN1   VNDL..
C_IN_93IN9   VNDL..
C_IN_93IN9   VNDM..
C_IN_94IN1   VNDL..
C_IN_95IN2   VNDL..
CRF01_AE_C   GDNL..
CRF01_AE_C   GDNL..
CRF01_AE_C   GDNV..
CRF01_AE_T   GDNL..
CRF01_AE_T   GDNL..
CRF01_AE_T   GDNV..
CRF01_AE_T   GDNL..
CRF01_AE_T   GDNL..
CRF01_AE_T   GDNL..
CRF02_AG_F   DNDNL.
CRF02_AG_F   DNDNL.
CRF02_AG_G   DNL...
CRF02_AG_N   DNDNL.
CRF02_AG_S   DNDNL.
CRF02_AG_S   DIDNL.
CRF03_AB_R   ADDL..
CRF03_AB_R   ADDL..
CRF04_cpx_   GDNL..
CRF04_cpx_   GNNV..
CRF04_cpx_   GNNL..
CRF05_DF_B   INNL..
CRF05_DF_B   INNL..
CRF06_cpx_   GDNL..
CRF06_cpx_   GDNL..
CRF06_cpx_   GDNL..
CRF06_cpx_   GDNL..
CRF11_cpx_   VNDL..
CRF11_cpx_   VNDL..
D_CD_84ZR0   VDDDL.
D_CD_ELI_K   IDDL..
D_CD_NDK_M   VDDL..
D_UG_94UG1   ADDM..
F1_BE_VI85   INNL..
F1_BR_93BR   INNL..
F1_FI_FIN9   VNNL..
F1_FR_MP41   INNL..
F2_CM_MP25   INNL..
F2KU_BE_VI   IHNL..
G_BE_DRCBL   GDNL..
G_NG_92NG0   GNNL..
G_SE_SE616   GDNL..
H_BE_VI991   VADL..
H_BE_VI997   VADL..
H_CF_90CF0   VADL..
J_SE_SE702   VNDL..
J_SE_SE788   VNDL..
K_CD_EQTB1   INNL..
K_CM_MP535   IDNL..
N_CM_YBF30   QGNWV.
O_CM_ANT70   NPMFEP
O_CM_MVP51   NPMFEL
```

```
O_SN_99SE_   NPMFEL
O_SN_99SE_   NPMFEP
U_CD___83C   DNDM..
```

Table 20. BLASTP Sequences producing significant alignments with S20757 (HBV Polymerase subtype ayw)

	Score (bits)	E: Value
gi|93080|pir||S20757 DNA-directed DNA polymerase (EC 2.7.7....	1553	0.0
gi|8925755|gb|AAF81607.1| DNA polymerase/reverse transcript...	1489	0.0
gi|1514497|emb|CAA68864.1| P [Hepatitis B virus]	1488	0.0
gi|27466573|gb|AAO12632.1| polymerase [Hepatitis B virus]	1482	0.0
gi|5257489|gb|AAD41360.1| polymerase [Hepatitis B virus]	1482	0.0
gi|118876|sp|P03156|DPOL_HPBVY P protein [Includes: DNA-dir...	1482	0.0
gi|27466565|gb|AAO12625.1| polymerase [Hepatitis B virus]	1481	0.0
gi|67003|pir||JDVLVB DNA-directed DNA polymerase (EC 2.7.7....	1480	0.0
gi|59433|emb|CAA46352.1| polymerase ORF [Hepatitis B virus]	1480	0.0
gi|6692498|gb|AAF24666.1| polymerase [Hepatitis B virus]	1479	0.0
gi|6692505|gb|AAF24673.1| polymerase [Hepatitis B virus]	1479	0.0
gi|2117935|pir||S71785 DNA-directed DNA polymerase (EC 2.7....	1477	0.0
gi|28436101|dbj|BAC57445.1| polymerase [Hepatitis B virus]	1476	0.0
gi|631984|pir||S47406 DNA-directed DNA polymerase (EC 2.7.7...	1475	0.0
gi|1359687|emb|CAA66431.1| polymerase [Hepatitis B virus]	1474	0.0
gi|18621117|emb|CAC87021.1| polymerase [Hepatitis B virus]	1474	0.0
gi|28436091|dbj|BAC57437.1| polymerase [Hepatitis B virus]	1473	0.0
gi|6692512|gb|AAF24680.1| polymerase [Hepatitis B virus]	1472	0.0
gi|22135695|gb|AAM09037.1| polymerase [Hepatitis B virus]	1471	0.0
gi|18621125|emb|CAC87015.1| polymerase [Hepatitis B virus]	1471	0.0
gi|1359679|emb|CAA66424.1| polymerase [Hepatitis B virus]	1470	0.0
gi|6692492|gb|AAF24660.1| polymerase [Hepatitis B virus]	1468	0.0
gi|2182121|gb|AAB59972.1| DNA polymerase [Hepatitis B virus]	1467	0.0
gi|4140295|emb|CAA10539.1| polymerase [Hepatitis B virus]	1467	0.0
gi|28436096|dbj|BAC57441.1| polymerase [Hepatitis B virus]	1466	0.0
gi|2829156|gb|AAC40810.1| polymerase [Hepatitis B virus]	1464	0.0
gi|27466519|gb|AAO12604.1| polymerase [Hepatitis B virus] >...	1463	0.0
gi|118869|sp|P24024|DPOL_HPBVA P protein [Includes: DNA-dir...	1462	0.0
gi|27466525|gb|AAO12672.1| polymerase [Hepatitis B virus]	1461	0.0
gi|762933|emb|CAA59514.1| polymerase [Hepatitis B virus]	1461	0.0
gi|22135690|gb|AAM09033.1| polymerase [Hepatitis B virus]	1459	0.0
gi|6063470|dbj|BAA85377.1| DNA polymerase/reverse transcrip...	1455	0.0
gi|6063465|dbj|BAA85373.1| DNA polymerase/reverse transcrip...	1454	0.0
gi|27466605|gb|AAO12660.1| polymerase [Hepatitis B virus]	1451	0.0
gi|2829149|gb|AAC40804.1| polymerase [Hepatitis B virus]	1451	0.0
gi|475987|gb|AAA18583.1| polymerase [Hepatitis B virus]	1450	0.0
gi|313784|emb|CAA42466.1| polymerase [Hepatitis B virus]	1446	0.0
gi|27466597|gb|AAO12653.1| polymerase [Hepatitis B virus]	1444	0.0
gi|15419833|gb|AAK97182.1|AF297620_3 polymerase [Hepatitis ...	1442	0.0
gi|93082|pir||S20752 DNA-directed DNA polymerase (EC 2.7.7....	1441	0.0
gi|27466613|gb|AAO12667.1| polymerase [Hepatitis B virus]	1435	0.0
gi|27466589|gb|AAO12646.1| polymerase [Hepatitis B virus]	1434	0.0
gi|27466538|gb|AAO12618.1| polymerase [Hepatitis B virus]	1432	0.0
gi|27466581|gb|AAO12639.1| polymerase [Hepatitis B virus]	1431	0.0
gi|15419828|gb|AAK97178.1|AF297619_3 polymerase [Hepatitis ...	1429	0.0
gi|27466544|gb|AAO12681.1| polymerase [Hepatitis B virus]	1427	0.0
gi|27466557|gb|AAO12692.1| polymerase [Hepatitis B virus]	1423	0.0
gi|16751312|gb|AAL25951.1| polymerase protein [Hepatitis B ...	1382	0.0
gi|11935073|gb|AAG41955.1|AF305327_2 polymerase [Hepatitis ...	1379	0.0
gi|13491150|gb|AAK27856.1|AF330110_3 polymerase [Hepatitis ...	1368	0.0
gi|6116700|dbj|BAA32859.2| pol protein [Hepatitis B virus]	1368	0.0
gi|3551332|dbj|BAA32886.1| pol protein [Hepatitis B virus]	1368	0.0
gi|28812222|dbj|BAC65108.1| polymerase protein [Hepatitis B...	1368	0.0
gi|6691505|dbj|BAA89330.1| polymerase protein [Hepatitis B ...	1368	0.0
gi|118872|sp|P12900|DPOL_HPBVL P protein [Includes: DNA-dir...	1368	0.0
gi|560084|dbj|BAA04927.1| DNA polymerase [Hepatitis B virus]	1367	0.0

gi|560089|dbj|BAA04931.1| DNA polymerase [Hepatitis B virus]	1367	0.0
gi|6116731|dbj|BAA32957.2| pol protein [Hepatitis B virus]	1366	0.0
gi|6691495|dbj|BAA89322.1| polymerase protein [Hepatitis B ...	1365	0.0
gi|7188655|gb|AAF37833.1|AF222323_2 polymerase [Hepatitis B...	1365	0.0
gi|6063460|dbj|BAA85369.1| DNA polymerase/reverse transcrip...	1364	0.0
gi|3551347|dbj|BAA32898.1| pol protein [Hepatitis B virus]	1364	0.0
gi|6691500|dbj|BAA89326.1| polymerase protein [Hepatitis B ...	1363	0.0
gi|28812217|dbj|BAC65104.1| polymerase protein [Hepatitis B...	1363	0.0
gi|3551342|dbj|BAA32894.1| pol protein [Hepatitis B virus]	1363	0.0
gi|628080|pir||S43491 DNA-directed DNA polymerase (EC 2.7.7...	1363	0.0
gi|12246972|gb|AAG49670.1|AF223956_3 polymerase [Hepatitis ...	1362	0.0
gi|3551293|dbj|BAA32852.1| pol protein [Hepatitis B virus]	1362	0.0
gi|12246964|gb|AAG49663.1|AF223955_3 polymerase [Hepatitis ...	1362	0.0
gi|21624231|dbj|BAC01103.1| polymerase protein [Hepatitis B...	1362	0.0
gi|118874|sp|P03157|DPOL_HPBVR P protein [Includes: DNA-dir...	1361	0.0
gi|6009784|dbj|BAA85065.1| polymerase [Hepatitis B virus]	1361	0.0
gi|22651880|gb|AAN03491.1|AF286594_3 DNA polymerase [Hepati...	1360	0.0
gi|18252591|gb|AAL66348.1|AF461043_2 P protein [Hepatitis B...	1360	0.0
gi|15778326|gb|AAL07381.1|AF411409_4 polymerase [Hepatitis ...	1360	0.0
gi|3551268|dbj|BAA32832.1| pol protein [Hepatitis B virus]	1360	0.0
gi|14290241|gb|AAK59316.1|AF384371_2 polymerase [Hepatitis ...	1358	0.0
gi|6063435|dbj|BAA85353.1| DNA polymerase/reverse transcrip...	1358	0.0
gi|6063440|dbj|BAA85357.1| DNA polymerase/reverse transcrip...	1358	0.0
gi|3551283|dbj|BAA32844.1| pol protein [Hepatitis B virus]	1358	0.0
gi|18252536|gb|AAL66307.1|AF458664_3 polymerase [Hepatitis ...	1358	0.0
gi|6009769|dbj|BAA85053.1| polymerase [Hepatitis B virus]	1358	0.0
gi|13991865|gb|AAK51533.1|AF363961_2 polymerase [Hepatitis ...	1357	0.0
gi|6063425|dbj|BAA85382.1| DNA polymerase/reverse transcrip...	1357	0.0
gi|2626986|dbj|BAA23435.1| DNA polymerase [Hepatitis B viru...	1357	0.0
gi|4490402|emb|CAB38767.1| P protein [Hepatitis B virus]	1357	0.0
gi|22415735|gb|AAM95242.1| DNA polymerase/reverse transcrip...	1357	0.0
gi|10934057|dbj|BAB16885.1| polymerase [Hepatitis B virus]	1356	0.0
gi|18252556|gb|AAL66323.1|AF461359_3 polymerase [Hepatitis ...	1356	0.0
gi|2627009|dbj|BAA23455.1| DNA polymerase [Hepatitis B virus]	1356	0.0
gi|560074|dbj|BAA04919.1| DNA polymerase [Hepatitis B virus]	1356	0.0
gi|479847|pir||S35527 DNA-directed DNA polymerase (EC 2.7.7...	1356	0.0
gi|18252545|gb|AAL66314.1|AF461357_2 polymerase [Hepatitis ...	1356	0.0
gi|1742906|dbj|BAA09083.1| DNA polymerase [Hepatitis B virus]	1355	0.0
gi|6009764|dbj|BAA85049.1| polymerase [Hepatitis B virus] >...	1355	0.0
gi|2627002|dbj|BAA23449.1| DNA polymerase [Hepatitis B virus]	1355	0.0
gi|6063455|dbj|BAA85365.1| DNA polymerase/reverse transcrip...	1355	0.0
gi|10441115|gb|AAG16953.1|AF182804_4 polymerase [Hepatitis ...	1354	0.0
gi|6009774|dbj|BAA85057.1| polymerase [Hepatitis B virus]	1353	0.0
gi|4490407|emb|CAB38771.1| P protein [Hepatitis B virus]	1353	0.0
gi|3582359|dbj|BAA32913.1| pol protein [Hepatitis B virus]	1353	0.0
gi|3582355|dbj|BAA32874.1| pol protein [Hepatitis B virus]	1353	0.0
gi|12246980|gb|AAG49677.1|AF223957_3 polymerase [Hepatitis ...	1352	0.0
gi|16751307|gb|AAL25947.1| polymerase protein [Hepatitis B ...	1352	0.0
gi|3582375|dbj|BAA32925.1| pol protein [Hepatitis B virus]	1352	0.0
gi|15778340|gb|AAL07392.1|AF411412_4 polymerase [Hepatitis ...	1352	0.0
gi|4206637|gb|AAD11755.1| DNA polymerase [Hepatitis B virus]	1352	0.0
gi|15425690|dbj|BAB64319.1| polymerase [Hepatitis B virus]	1352	0.0
gi|3551352|dbj|BAA32902.1| pol protein [Hepatitis B virus]	1352	0.0
gi|3582395|dbj|BAA32963.1| pol protein [Hepatitis B virus]	1352	0.0
gi|5114071|gb|AAD40205.1|AF090839_2 polymerase [Hepatitis B...	1352	0.0
gi|9082085|gb|AAF82723.1|AF233236_2 pol [Hepatitis B virus]	1352	0.0
gi|6983935|gb|AAF34734.1|AF160501_2 polymerase [Hepatitis B...	1351	0.0
gi|560094|dbj|BAA04935.1| DNA polymerase [Hepatitis B virus]	1351	0.0
gi|18032033|gb|AAL49990.1| polymerase [Hepatitis B virus]	1351	0.0
gi|18146671|dbj|BAB82392.1| polymerase [Hepatitis B virus]	1351	0.0
gi|6006322|dbj|BAA84819.1| polymerase protein [Hepatitis B ...	1350	0.0

gi|18252551|gb|AAL66319.1|AF461358_3 polymerase [Hepatitis ...	1350	0.0
gi|7188649|gb|AAF37828.1|AF222322_2 polymerase [Hepatitis B...	1350	0.0
gi|12060441|dbj|BAB20611.1| DNA polymerase [Hepatitis B virus]	1350	0.0
gi|18845085|gb|AAL79545.1|AF473543_4 P protein [Hepatitis B...	1350	0.0
gi|3551322|dbj|BAA32878.1| pol protein [Hepatitis B virus]	1350	0.0
gi|12246956|gb|AAG49656.1|AF223954_4 polymerase [Hepatitis ...	1350	0.0
gi|6063430|dbj|BAA85349.1| DNA polymerase/reverse transcrip...	1350	0.0
gi|2288872|dbj|BAA21665.1| DNA polymerase [Hepatitis B virus]	1350	0.0
gi|1220111|dbj|BAA04072.1| DNA polymerase [Hepatitis B virus]	1349	0.0
gi|9454168|gb|AAF87689.1| polymerase protein [Hepatitis B v...	1349	0.0
gi|18146683|dbj|BAB82402.1| polymerase [Hepatitis B virus]	1349	0.0
gi|3551278|dbj|BAA32840.1| pol protein [Hepatitis B virus]	1349	0.0
gi|3551372|dbj|BAA32939.1| pol protein [Hepatitis B virus]	1349	0.0
gi|19849035|gb|AAL99437.1|AF405706_3 polymerase [Hepatitis ...	1349	0.0
gi|3551357|dbj|BAA32906.1| pol protein [Hepatitis B virus]	1349	0.0
gi|15778321|gb|AAL07377.1|AF411408_4 polymerase [Hepatitis ...	1348	0.0
gi|15072542|gb|AAK81690.1| polymerase protein [Hepatitis B ...	1348	0.0
gi|21624238|dbj|BAC01109.1| polymerase protein [Hepatitis B...	1348	0.0
gi|12247012|gb|AAG49705.1|AF223961_3 polymerase [Hepatitis ...	1348	0.0
gi|5114086|gb|AAD40217.1|AF090842_2 polymerase [Hepatitis B...	1348	0.0
gi|3582407|dbj|BAA32972.1| pol protein [Hepatitis B virus]	1347	0.0
gi|15425698|dbj|BAB64325.1| polymerase [Hepatitis B virus]	1347	0.0
gi|18146665|dbj|BAB82387.1| polymerase [Hepatitis B virus]	1347	0.0
gi|23194252|gb|AAN15074.1| P protein [Hepatitis B virus]	1347	0.0
gi|560079|dbj|BAA04923.1| DNA polymerase [Hepatitis B virus]	1347	0.0
gi|10443833|gb|AAG17595.1|AF241410_3 polymerase [Hepatitis ...	1346	0.0
gi|13991870|gb|AAK51537.1|AF363962_2 polymerase [Hepatitis ...	1346	0.0
gi|4007054|emb|CAA10426.1| DNA polymerase [Hepatitis B virus]	1346	0.0
gi|3551362|dbj|BAA32910.1| pol protein [Hepatitis B virus]	1346	0.0
gi|18146677|dbj|BAB82397.1| polymerase [Hepatitis B virus]	1346	0.0
gi|12246988|gb|AAG49684.1|AF223958_3 polymerase [Hepatitis ...	1346	0.0
gi|15211897|emb|CAC51286.1| polymerase [Hepatitis B virus]	1345	0.0
gi|18389989|gb|AAL68823.1| polymerase [Hepatitis B virus]	1345	0.0
gi|3582363|dbj|BAA32916.1| pol protein [Hepatitis B virus]	1345	0.0
gi|10441110|gb|AAG16949.1|AF182803_4 polymerase [Hepatitis ...	1345	0.0
gi|10443841|gb|AAG17602.1|AF241411_3 polymerase [Hepatitis ...	1345	0.0
gi|3551382|dbj|BAA32947.1| pol protein [Hepatitis B virus]	1345	0.0
gi|3582387|dbj|BAA32950.1| pol protein [Hepatitis B virus]	1344	0.0
gi|3551317|dbj|BAA32871.1| pol protein [Hepatitis B virus]	1344	0.0
gi|10441104|gb|AAG16944.1|AF182802_3 polymerase [Hepatitis ...	1343	0.0
gi|118866|sp|P03159|DPOL_HPBV2 P protein [Includes: DNA-dir...	1343	0.0
gi|15425694|dbj|BAB64322.1| polymerase [Hepatitis B virus]	1343	0.0
gi|4007049|emb|CAA10422.1| DNA polymerase [Hepatitis B virus]	1343	0.0
gi|29123239|gb|AAO62971.1| pol protein [Hepatitis B virus]	1343	0.0
gi|4007064|emb|CAA10438.1| DNA polymerase [Hepatitis B virus]	1342	0.0
gi|452623|emb|CAA53358.1| polymerase [Hepatitis B virus]	1342	0.0
gi|18252541|gb|AAL66311.1|AF458665_3 polymerase [Hepatitis ...	1342	0.0
gi|527443|emb|CAA84791.1| DNA polymerase [Hepatitis B virus]	1342	0.0
gi|15211890|emb|CAC51280.1| polymerase [Hepatitis B virus]	1342	0.0
gi|329617|gb|AAA62812.1| DNA polymerase	1341	0.0
gi|4007079|emb|CAA10454.1| DNA polymerase [Hepatitis B virus]	1341	0.0
gi|9454173|gb|AAF87693.1| polymerase protein [Hepatitis B v...	1341	0.0
gi|452628|emb|CAA53354.1| polymerase [Hepatitis B virus]	1341	0.0
gi|3582367|dbj|BAA32919.1| pol protein [Hepatitis B virus]	1340	0.0
gi|5114066|gb|AAD40201.1|AF090838_2 polymerase [Hepatitis B...	1340	0.0
gi|15419860|gb|AAK97203.1|AF297625_3 polymerase [Hepatitis ...	1340	0.0
gi|4490412|emb|CAB38775.1| P protein [Hepatitis B virus]	1340	0.0
gi|18252566|gb|AAL66331.1|AF461361_3 polymerase [Hepatitis ...	1340	0.0
gi|4007059|emb|CAA10430.1| DNA polymerase [Hepatitis B virus]	1340	0.0
gi|5114081|gb|AAD40213.1|AF090841_2 polymerase [Hepatitis B...	1339	0.0
gi|3582371|dbj|BAA32922.1| pol protein [Hepatitis B virus]	1339	0.0

gi|12247003|gb|AAG49697.1|AF223960_4 polymerase [Hepatitis ...	1339	0.0
gi|4033548|emb|CAA10450.1| DNA polymerase [Hepatitis B virus]	1339	0.0
gi|3892581|emb|CAA09962.1| polymerase [Hepatitis B virus]	1339	0.0
gi|5114076|gb|AAD40209.1|AF090840_2 polymerase [Hepatitis B...	1338	0.0
gi|12060436|dbj|BAB20607.1| DNA polymerase [Hepatitis B virus]	1338	0.0
gi|118868|sp|P17100|DPOL_HPBV9 P protein [Includes: DNA-dir...	1337	0.0
gi|27466434|gb|AAO12555.1| polymerase [Hepatitis B virus]	1337	0.0
gi|3582399|dbj|BAA32966.1| pol protein [Hepatitis B virus]	1337	0.0
gi|3551273|dbj|BAA32836.1| pol protein [Hepatitis B virus]	1337	0.0
gi|14285168|gb|AAK58873.1| polymerase [synthetic construct]...	1337	0.0
gi|3582391|dbj|BAA32953.1| pol protein [Hepatitis B virus]	1337	0.0
gi|15419845|gb|AAK97191.1|AF297622_3 polymerase [Hepatitis ...	1337	0.0
gi|118870|sp|P17393|DPOL_HPBVI P protein [Includes: DNA-dir...	1336	0.0
gi|3551377|dbj|BAA32943.1| pol protein [Hepatitis B virus]	1336	0.0
gi|10443825|gb|AAG17588.1|AF241409_3 polymerase [Hepatitis ...	1336	0.0
gi|10443817|gb|AAG17581.1|AF241408_3 polymerase [Hepatitis ...	1336	0.0
gi|29124889|gb|AAO63519.1| pol protein [Hepatitis B virus]	1335	0.0
gi|399401|sp|P31870|DPOL_HPBVM P protein [Includes: DNA-dir...	1335	0.0
gi|6063445|dbj|BAA85339.1| DNA polymerase/reverse transcrip...	1335	0.0
gi|19568078|gb|AAL89566.1| polymerase [Hepatitis B virus]	1334	0.0
gi|27466426|gb|AAO12548.1| polymerase [Hepatitis B virus]	1334	0.0
gi|22655601|gb|AAN04128.1| polymerase [Hepatitis B virus]	1334	0.0
gi|8161369|gb|AAA69721.2| polymerase [Hepatitis B virus]	1334	0.0
gi|10441120|gb|AAG16957.1|AF182805_4 polymerase [Hepatitis ...	1334	0.0
gi|10443809|gb|AAG17574.1|AF241407_3 polymerase [Hepatitis ...	1333	0.0
gi|18146689|dbj|BAB82407.1| polymerase [Hepatitis B virus]	1333	0.0
gi|4007069|emb|CAA10442.1| DNA polymerase [Hepatitis B virus]	1333	0.0
gi|18031709|gb|AAK57744.1| polymerase [Hepatitis B virus]	1333	0.0
gi|18252561|gb|AAL66327.1|AF461360_3 polymerase [Hepatitis ...	1332	0.0
gi|6959503|gb|AAF33121.1| polymerase protein [orangutan hep...	1332	0.0
gi|26224721|gb|AAN76318.1| polymerase [Hepatitis B virus]	1332	0.0
gi|4007074|emb|CAA10446.1| DNA polymerase [Hepatitis B virus]	1332	0.0
gi|18031714|gb|AAK57745.1| polymerase [Hepatitis B virus]	1332	0.0
gi|7434791|pir||S67505 DNA-directed DNA polymerase (EC 2.7....	1332	0.0
gi|15419855|gb|AAK97199.1|AF297624_3 polymerase [Hepatitis ...	1332	0.0
gi|7434793|pir||T13468 DNA-directed DNA polymerase (EC 2.7....	1331	0.0
gi|4323205|gb|AAD16257.1| polymerase [Hepatitis B virus]	1331	0.0
gi|12060194|dbj|BAB20451.1| DNA polymerase [Hepatitis B virus]	1331	0.0
gi|23194347|gb|AAN15122.1| polymerase [Hepatitis B virus]	1330	0.0
gi|20151228|gb|AAM12945.1| DNA polymerase/reverse transcrip...	1330	0.0
gi|23884547|gb|AAN40009.1| pol protein [Hepatitis B virus]	1330	0.0
gi|21431681|gb|AAM53414.1|U87747_3 DNA polymerase/reverse t...	1330	0.0
gi|3551337|dbj|BAA32890.1| pol protein [Hepatitis B virus]	1329	0.0
gi|5019933|gb|AAD37919.1| P protein [Hepatitis B virus]	1329	0.0
gi|15419840|gb|AAK97187.1|AF297621_3 polymerase [Hepatitis ...	1329	0.0
gi|6006331|dbj|BAA84825.1| polymerase protein [Hepatitis B ...	1329	0.0
gi|19568073|gb|AAL89569.1| polymerase [Hepatitis B virus]	1329	0.0
gi|29124918|gb|AAO63539.1| pol protein [Hepatitis B virus]	1328	0.0
gi|329630|gb|AAA45483.1| P protein [Hepatitis B virus]	1328	0.0
gi|15778331|gb|AAL07385.1|AF411410_4 polymerase [Hepatitis ...	1328	0.0
gi|6566410|dbj|BAA88275.1| P protein [Hepatitis B virus]	1328	0.0
gi|4490397|emb|CAB38763.1| P protein [Hepatitis B virus]	1328	0.0
gi|12060187|dbj|BAB20445.1| DNA polymerase [Hepatitis B virus]	1327	0.0
gi|6063450|dbj|BAA85343.1| DNA polymerase/reverse transcrip...	1327	0.0
gi|118877|sp|P03155|DPOL_HPBVZ P protein [Includes: DNA-dir...	1327	0.0
gi|29124883|gb|AAO63514.1| pol protein [Hepatitis B virus]	1325	0.0
gi|4033543|emb|CAA10434.1| DNA polymerase [Hepatitis B virus]	1325	0.0
gi|6692525|gb|AAF24693.1| polymerase [Hepatitis B virus]	1325	0.0
gi|6692559|gb|AAF24727.1| polymerase [Hepatitis B virus]	1325	0.0
gi|23194340|gb|AAN15116.1| polymerase [Hepatitis B virus]	1325	0.0
gi|560064|dbj|BAA04911.1| DNA polymerase [Hepatitis B virus]	1324	0.0

gi|29124898|gb|AAO63526.1|	pol protein [Hepatitis B virus]	1324	0.0
gi|29124927|gb|AAO63545.1|	pol protein [Hepatitis B virus]	1323	0.0
gi|6692566|gb|AAF24734.1|	polymerase [Hepatitis B virus]	1323	0.0
gi|6692553|gb|AAF24721.1|	polymerase [Hepatitis B virus]	1323	0.0
gi|6692518|gb|AAF24686.1|	polymerase [Hepatitis B virus] >g...	1323	0.0
gi|1359702|emb|CAA66444.1|	polymerase [Hepatitis B virus] >...	1323	0.0
gi|29124867|gb|AAO63501.1|	pol protein [Hepatitis B virus]	1323	0.0
gi|29124872|gb|AAO63505.1|	pol protein [Hepatitis B virus] ...	1323	0.0
gi|27466479|gb|AAO12576.1|	polymerase [Hepatitis B virus]	1322	0.0
gi|6692546|gb|AAF24714.1|	polymerase [Hepatitis B virus]	1322	0.0
gi|3551312|dbj|BAA32867.1|	pol protein [Hepatitis B virus]	1322	0.0
gi|27466487|gb|AAO12611.1|	polymerase [Hepatitis B virus]	1322	0.0
gi|118871|sp|P17394|DPOL_HPBVJ	P protein [Includes: DNA-dir...	1321	0.0
gi|9454473|gb|AAF87833.1|AF282917_3	DNA polymerse [Hepatiti...	1321	0.0
gi|19224214|gb|AAL86445.1|AF479684_3	P gene product [Hepati...	1321	0.0
gi|6692572|gb|AAF24740.1|	polymerase [Hepatitis B virus]	1321	0.0
gi|3551297|dbj|BAA32855.1|	pol protein [Hepatitis B virus]	1321	0.0
gi|3551327|dbj|BAA32882.1|	pol protein [Hepatitis B virus]	1320	0.0
gi|1359695|emb|CAA66434.1|	polymerase [Hepatitis B virus]	1320	0.0
gi|3551367|dbj|BAA32932.1|	pol protein [Hepatitis B virus]	1320	0.0
gi|118873|sp|P17395|DPOL_HPBVO	P protein [Includes: DNA-dir...	1319	0.0
gi|29124862|gb|AAO63497.1|	pol protein [Hepatitis B virus]	1319	0.0
gi|18621110|emb|CAC87028.1|	polymerase [Hepatitis B virus]	1319	0.0
gi|3582403|dbj|BAA32969.1|	pol protein [Hepatitis B virus]	1318	0.0
gi|27261550|gb|AAN85925.1|	DNA polymerase [Hepatitis B viru...	1318	0.0
gi|1914703|emb|CAA66699.1|	polymerase [Hepatitis B virus]	1318	0.0
gi|4323200|gb|AAD16253.1|	polymerase [Hepatitis B virus]	1318	0.0
gi|6573293|dbj|BAA88291.1|	P protein [Hepatitis B virus]	1318	0.0
gi|6006341|dbj|BAA84833.1|	polymerase protein [Hepatitis B ...	1316	0.0
gi|6566440|dbj|BAA88286.1|	P protein [Hepatitis B virus]	1315	0.0
gi|560059|dbj|BAA04907.1|	DNA polymerase [Hepatitis B virus]	1315	0.0
gi|14334410|gb|AAK59391.1|	polymerase [Hepatitis B virus]	1315	0.0
gi|5019954|gb|AAD37936.1|	P protein [Hepatitis B virus]	1315	0.0
gi|16117323|dbj|BAB69785.1|	polymerase [Hepatitis B virus]	1315	0.0
gi|7434792|pir||T13473	DNA-directed DNA polymerase (EC 2.7....	1315	0.0
gi|5019965|gb|AAD37945.1|	P protein [Hepatitis B virus]	1314	0.0
gi|29124908|gb|AAO63533.1|	pol protein [Hepatitis B virus]	1314	0.0
gi|6566428|dbj|BAA88281.1|	P protein [Hepatitis B virus]	1313	0.0
gi|29124894|gb|AAO63523.1|	pol protein [Hepatitis B virus]	1311	0.0
gi|22135730|gb|AAM09065.1|	polymerase [Hepatitis B virus]	1311	0.0
gi|560069|dbj|BAA04915.1|	DNA polymerase [Hepatitis B virus]	1311	0.0
gi|15419850|gb|AAK97195.1|AF297623_3	polymerase [Hepatitis ...	1311	0.0
gi|9634217|ref|NP_037757.1|	polymerase protein [orangutan h...	1310	0.0
gi|16117333|dbj|BAB69793.1|	polymerase [Hepatitis B virus]	1309	0.0
gi|9971630|dbj|BAB12582.1|	polymerase protein [Hepatitis B ...	1308	0.0
gi|27466450|gb|AAO12569.1|	polymerase [Hepatitis B virus]	1306	0.0
gi|12247036|gb|AAG49726.1|AF223964_3	polymerase [Hepatitis ...	1306	0.0
gi|12247028|gb|AAG49719.1|AF223963_3	polymerase [Hepatitis ...	1305	0.0
gi|5019945|gb|AAD37929.1|	P protein [Hepatitis B virus]	1305	0.0
gi|18146701|dbj|BAB82417.1|	polymerase [Hepatitis B virus]	1305	0.0
gi|12247020|gb|AAG49712.1|AF223962_3	polymerase [Hepatitis ...	1304	0.0
gi|5019981|gb|AAD37958.1|	P protein [Hepatitis B virus]	1304	0.0
gi|3892582|emb|CAA53343.1|	polymerase [Hepatitis B virus]	1304	0.0
gi|27466442|gb|AAO12562.1|	polymerase [Hepatitis B virus]	1304	0.0
gi|22135715|gb|AAM09053.1|	polymerase [Hepatitis B virus]	1301	0.0
gi|12247044|gb|AAG49733.1|AF223965_3	polymerase [Hepatitis ...	1301	0.0
gi|22135725|gb|AAM09061.1|	polymerase [Hepatitis B virus]	1301	0.0
gi|11191880|dbj|BAB17962.1|	polymerase [Hepatitis B virus]	1300	0.0
gi|3551392|dbj|BAA32961.1|	pol protein [Hepatitis B virus]	1300	0.0
gi|6006336|dbj|BAA84829.1|	polymerase protein [Hepatitis B ...	1299	0.0
gi|2627021|dbj|BAA23467.1|	DNA polymerase [Hepatitis B virus]	1298	0.0

gi|2627015|dbj|BAA23461.1| DNA polymerase [Hepatitis B virus]	1297	0.0
gi|16117328|dbj|BAB69789.1| polymerase [Hepatitis B virus]	1297	0.0
gi|22135735|gb|AAM09069.1| polymerase [Hepatitis B virus]	1297	0.0
gi|14485226|gb|AAK62976.1|AF384372_2 polymerase [Hepatitis ...	1296	0.0
gi|3551288|dbj|BAA32848.1| pol protein [Hepatitis B virus]	1295	0.0
gi|11191960|dbj|BAB18032.1| polymerase [Hepatitis B virus]	1294	0.0
gi|11191888|dbj|BAB17969.1| polymerase [Hepatitis B virus] ...	1293	0.0
gi|11191840|dbj|BAB17927.1| polymerase [Hepatitis B virus] ...	1293	0.0
gi|11191920|dbj|BAB17997.1| polymerase [Hepatitis B virus]	1293	0.0
gi|11191904|dbj|BAB17983.1| polymerase [Hepatitis B virus]	1291	0.0
gi|11191952|dbj|BAB18025.1| polymerase [Hepatitis B virus]	1291	0.0
gi|1169410|sp|Q05486|DPOL_HPBVT P protein [Includes: DNA-di...	1289	0.0
gi|22135705|gb|AAM09045.1| polymerase [Hepatitis B virus]	1288	0.0
gi|452633|emb|CAA53350.1| polymerase [Hepatitis B virus]	1288	0.0
gi|18146695|dbj|BAB82412.1| polymerase [Hepatitis B virus]	1287	0.0
gi|22135710|gb|AAM09049.1| polymerase [Hepatitis B virus]	1287	0.0
gi|11191864|dbj|BAB17948.1| polymerase [Hepatitis B virus]	1286	0.0
gi|59451|emb|CAA48354.1| HBV polymerase [Hepatitis B virus]	1286	0.0
gi|11191848|dbj|BAB17934.1| polymerase [Hepatitis B virus] ...	1286	0.0
gi|22135700|gb|AAM09041.1| polymerase [Hepatitis B virus]	1285	0.0
gi|5019976|gb|AAD37954.1| P protein [Hepatitis B virus]	1281	0.0
gi|22135720|gb|AAM09057.1| polymerase [Hepatitis B virus]	1279	0.0
gi|5019939|gb|AAD37924.1| P protein [Hepatitis B virus]	1276	0.0
gi|1914697|emb|CAA66674.1| polymerase [Hepatitis B virus]	1273	0.0
gi|1914691|emb|CAA66679.1| polymerase [Hepatitis B virus]	1271	0.0
gi|5019970|gb|AAD37949.1| P protein [Hepatitis B virus]	1263	0.0
gi|15425702|dbj|BAB64328.1| polymerase [Hepatitis B virus]	1258	0.0
gi|29124905|gb|AAO63531.1| pol protein [Hepatitis B virus]	1253	0.0
gi|27466464|gb|AAO12704.1| polymerase [Hepatitis B virus]	1248	0.0
gi|27466471|gb|AAO12710.1| polymerase [Hepatitis B virus]	1244	0.0
gi|18252571|gb|AAL66335.1|AF461362_3 polymerase [Hepatitis ...	1243	0.0
gi|27466511|gb|AAO12597.1| polymerase [Hepatitis B virus]	1239	0.0
gi|27466457|gb|AAO12698.1| polymerase [Hepatitis B virus]	1238	0.0
gi|15211905|emb|CAC51293.1| polymerase [Hepatitis B virus]	1227	0.0
gi|399402|sp|Q02314|DPOL_HPBVP P protein [Includes: DNA-dir...	1224	0.0
gi|1914708|emb|CAA66684.1| polymerase [Hepatitis B virus]	1220	0.0
gi|27466503|gb|AAO12583.1| polymerase [Hepatitis B virus]	1184	0.0
gi|118867|sp|P12933|DPOL_HPBV4 P protein [Includes: DNA-dir...	1157	0.0
gi|4468850|emb|CAB38229.1| polymerase [Hepatitis B virus]	1122	0.0
gi|1914719|emb|CAA66694.1| polymerase [Hepatitis B virus]	1101	0.0
gi|9630375|ref|NP_046799.1| polymerase [woolly monkey hepat...	1049	0.0
gi|1185115|emb|CAA51254.1| DNA polymerase [Hepatitis B virus]	922	0.0
gi|20800461|gb|AAM28642.1|U87746_4 DNA polymerase/reverse t...	910	0.0
gi|21326585|ref|NP_647604.1| P gene product (AA 304-843); c...	907	0.0
gi|4377612|emb|CAA53339.1| polymerase [Hepatitis B virus]	904	0.0
gi|4377613|emb|CAA53338.1| polymerase [Hepatitis B virus]	901	0.0
gi|1549226|dbj|BAA04073.1| ORF [Hepatitis B virus]	898	0.0
gi|9454414|gb|AAF87797.1| polymerase [Hepatitis B virus]	895	0.0
gi|1550614|dbj|BAA04075.1| ORF [Hepatitis B virus]	893	0.0
gi|59409|emb|CAA32399.1| DNA polymerase [Hepatitis B virus]	879	0.0
gi|118894|sp|P03160|DPOL_WHV1 P protein [Includes: DNA-dire...	727	0.0
gi|9626716|ref|NP_040994.1| A protein [Ground squirrel hepa...	727	0.0
gi|22256032|ref|NP_671813.1| DNA polymerase [Woodchuck hepa...	725	0.0
gi|15637595|gb|AAL04547.1|AF410859_1 polymerase [Woodchuck ...	725	0.0
gi|15637587|gb|AAL04543.1|AF410855_1 type II mutant polymer...	725	0.0
gi|118895|sp|P12899|DPOL_WHV59 P protein [Includes: DNA-dir...	724	0.0
gi|15637597|gb|AAL04548.1|AF410860_1 polymerase [Woodchuck ...	724	0.0
gi|15637599|gb|AAL04549.1|AF410861_1 polymerase [Woodchuck ...	722	0.0
gi|15637593|gb|AAL04546.1|AF410858_1 defective polymerase [...	721	0.0
gi|118898|sp|P17396|DPOL_WHV8I P protein [Includes: DNA-dir...	721	0.0
gi|15637591|gb|AAL04545.1|AF410857_1 type I mutant polymera...	721	0.0

gi|15637589|gb|AAL04544.1|AF410856_1 type IV mutant polymer...	717	0.0
gi|118897|sp|P06275|DPOL_WHV8 P protein [Includes: DNA-dire...	706	0.0
gi|3582379|dbj|BAA32928.1| pol protein [Hepatitis B virus]	692	0.0
gi|9885813|gb|AAG01539.1|AF291830_2 polymerase [Hepatitis B...	692	0.0
gi|118875|sp|P03158|DPOL_HPBVW DNA polymerase	680	0.0
gi|9628830|ref|NP_043864.1| polymerase [Arctic ground squir...	669	0.0
gi|8926931|dbj|BAA98025.1| pol protein [Hepatitis B virus]	669	0.0
gi|8926928|dbj|BAA98023.1| pol protein [Hepatitis B virus]	667	0.0
gi|8926925|dbj|BAA98021.1| pol protein [Hepatitis B virus]	667	0.0
gi|8926934|dbj|BAA98027.1| pol protein [Hepatitis B virus]	655	0.0
gi|13345982|gb|AAK19538.1|AF335734_2 polymerase [Hepatitis ...	583	e-166
gi|12083172|gb|AAG48743.1|AF329861_2 polymerase [Hepatitis ...	583	e-166
gi|13345979|gb|AAK19536.1|AF335733_2 polymerase [Hepatitis ...	583	e-166
gi|12083181|gb|AAG48749.1|AF329864_2 polymerase [Hepatitis ...	582	e-166
gi|12083178|gb|AAG48747.1|AF329863_2 polymerase [Hepatitis ...	582	e-165
gi|12083163|gb|AAG48737.1|AF329858_1 polymerase [Hepatitis ...	581	e-165
gi|12083167|gb|AAG48740.1|AF329859_2 polymerase [Hepatitis ...	581	e-165
gi|13345988|gb|AAK19542.1|AF335736_2 polymerase [Hepatitis ...	580	e-165
gi|13345985|gb|AAK19540.1|AF335735_2 polymerase [Hepatitis ...	578	e-164
gi|2982339|gb|AAC06361.1| DNA polymerase [Hepatitis B virus]	568	e-161
gi|336159|gb|AAA46774.1| polymerase protein	566	e-161
gi|118899|sp|P11292|DPOL_WHVW6 P protein [Includes: DNA-dir...	560	e-159
gi|225532|prf||1305266C gene P	555	e-157
gi|1107586|emb|CAA56892.1| polymerase [Hepatitis B virus]	540	e-153
gi|1107579|emb|CAA56878.1| polymerase [Hepatitis B virus]	538	e-152
gi|1185116|emb|CAA51255.1| HBsAg [Hepatitis B virus]	465	e-130
gi|59414|emb|CAA32405.1| DNA polymerase [Hepatitis B virus]	459	e-129
gi|21326589|ref|NP_647608.1| P gene product, put.DNA polyme...	458	e-128
gi|1321828|emb|CAA96556.1| polymerase [Hepatitis B virus]	441	e-123
gi|5019960|gb|AAD37941.1| P protein [Hepatitis B virus]	440	e-123
gi|329652|gb|AAA69719.1| coat protein [Hepatitis B virus]	433	e-121
gi|329651|gb|AAA69720.1| coat protein [Hepatitis B virus]	429	e-120
gi|27466495|gb|AAO12590.1| polymerase [Hepatitis B virus]	429	e-120
gi|21218028|dbj|BAB96528.1| large S protein [Hepatitis B vi...	413	e-115
gi|1321832|emb|CAA96561.1| polymerase [Hepatitis B virus]	410	e-114
gi|27450190|gb|AAO14552.1|AF460225_1 polymerase [Hepatitis ...	385	e-106
gi|27450188|gb|AAO14551.1|AF460224_1 polymerase [Hepatitis ...	384	e-106
gi|27450192|gb|AAO14553.1|AF460226_1 polymerase [Hepatitis ...	383	e-106
gi|27450198|gb|AAO14556.1|AF460229_1 polymerase [Hepatitis ...	382	e-105
gi|27450196|gb|AAO14555.1|AF460228_1 polymerase [Hepatitis ...	382	e-105
gi|27450194|gb|AAO14554.1|AF460227_1 polymerase [Hepatitis ...	382	e-105
gi|27450200|gb|AAO14557.1|AF460230_1 polymerase [Hepatitis ...	375	e-103
gi|27450202|gb|AAO14558.1|AF460231_1 polymerase [Hepatitis ...	375	e-103
gi|3328370|gb|AAC26832.1| DNA polymerase [Hepatitis B virus]	374	e-103
gi|23380174|gb|AAM83022.1| polymerase [Hepatitis B virus]	373	e-103
gi|23380081|gb|AAM82960.1| polymerase [Hepatitis B virus]	373	e-103
gi|23380171|gb|AAM83020.1| polymerase [Hepatitis B virus]	372	e-102
gi|23380180|gb|AAM83026.1| polymerase [Hepatitis B virus]	370	e-102
gi|23380177|gb|AAM83024.1| polymerase [Hepatitis B virus]	369	e-102
gi|23380072|gb|AAM82954.1| polymerase [Hepatitis B virus]	369	e-101
gi|23380084|gb|AAM82962.1| polymerase [Hepatitis B virus] >...	368	e-101
gi|23380078|gb|AAM82958.1| polymerase [Hepatitis B virus]	368	e-101
gi|23380066|gb|AAM82950.1| polymerase [Hepatitis B virus]	368	e-101
gi|23380111|gb|AAM82980.1| polymerase [Hepatitis B virus]	368	e-101
gi|23380063|gb|AAM82948.1| polymerase [Hepatitis B virus]	367	e-101
gi|23380087|gb|AAM82964.1| polymerase [Hepatitis B virus]	367	e-101
gi|3335627|gb|AAD13662.1| DNA polymerase [Hepatitis B virus]	366	e-101
gi|23380069|gb|AAM82952.1| polymerase [Hepatitis B virus]	366	e-101
gi|23380090|gb|AAM82966.1| polymerase [Hepatitis B virus]	366	e-101
gi|23380060|gb|AAM82946.1| polymerase [Hepatitis B virus]	366	e-101
gi|23380105|gb|AAM82976.1| polymerase [Hepatitis B virus]	365	e-100

gi|23380132|gb|AAM82994.1|	polymerase [Hepatitis B virus]		365	e-100
gi|23380093|gb|AAM82968.1|	polymerase [Hepatitis B virus]		365	e-100
gi|23380183|gb|AAM83028.1|	polymerase [Hepatitis B virus]	>...	365	e-100
gi|23380120|gb|AAM82986.1|	polymerase [Hepatitis B virus]		365	e-100
gi|13991875|gb|AAK51541.1|	AF363963_2 truncated polymerase	[...	365	e-100
gi|23380129|gb|AAM82992.1|	polymerase [Hepatitis B virus]		363	e-100
gi|23380186|gb|AAM83030.1|	polymerase [Hepatitis B virus]		363	e-100
gi|23380168|gb|AAM83018.1|	polymerase [Hepatitis B virus]		363	e-100
gi|23380075|gb|AAM82956.1|	polymerase [Hepatitis B virus]		363	e-100
gi|23380123|gb|AAM82988.1|	polymerase [Hepatitis B virus]		361	3e-99
gi|23380135|gb|AAM82996.1|	polymerase [Hepatitis B virus]		357	4e-98
gi|23380030|gb|AAM82926.1|	polymerase [Hepatitis B virus]		351	3e-96
gi|23380021|gb|AAM82920.1|	polymerase [Hepatitis B virus]		351	3e-96
gi|23379934|gb|AAM82862.1|	polymerase [Hepatitis B virus]	>...	350	8e-96
gi|23380036|gb|AAM82930.1|	polymerase [Hepatitis B virus]		349	1e-95
gi|23380156|gb|AAM83010.1|	polymerase [Hepatitis B virus]		349	1e-95
gi|23379922|gb|AAM82854.1|	polymerase [Hepatitis B virus]		349	1e-95
gi|23379943|gb|AAM82868.1|	polymerase [Hepatitis B virus]		349	1e-95
gi|23379967|gb|AAM82884.1|	polymerase [Hepatitis B virus]		349	1e-95
gi|23379928|gb|AAM82858.1|	polymerase [Hepatitis B virus]		349	1e-95
gi|23380057|gb|AAM82944.1|	polymerase [Hepatitis B virus]		348	2e-95
gi|23379925|gb|AAM82856.1|	polymerase [Hepatitis B virus]	>...	348	2e-95
gi|23380141|gb|AAM83000.1|	polymerase [Hepatitis B virus]		348	2e-95
gi|23380165|gb|AAM83016.1|	polymerase [Hepatitis B virus]		348	2e-95
gi|23379997|gb|AAM82904.1|	polymerase [Hepatitis B virus]		348	3e-95
gi|23380147|gb|AAM83004.1|	polymerase [Hepatitis B virus]	>...	348	3e-95
gi|23379868|gb|AAM82818.1|	polymerase [Hepatitis B virus]	>...	348	3e-95
gi|23379958|gb|AAM82878.1|	polymerase [Hepatitis B virus]	>...	348	3e-95
gi|23379904|gb|AAM82842.1|	polymerase [Hepatitis B virus]		347	3e-95
gi|23379931|gb|AAM82860.1|	polymerase [Hepatitis B virus]		347	3e-95
gi|23380159|gb|AAM83012.1|	polymerase [Hepatitis B virus]		347	3e-95
gi|23380144|gb|AAM83002.1|	polymerase [Hepatitis B virus]		347	3e-95
gi|23379892|gb|AAM82834.1|	polymerase [Hepatitis B virus]		347	4e-95
gi|23380000|gb|AAM82906.1|	polymerase [Hepatitis B virus]		347	4e-95
gi|23380042|gb|AAM82934.1|	polymerase [Hepatitis B virus]		347	4e-95
gi|23380003|gb|AAM82908.1|	polymerase [Hepatitis B virus]		347	5e-95
gi|23379886|gb|AAM82830.1|	polymerase [Hepatitis B virus]		347	5e-95
gi|23380009|gb|AAM82912.1|	polymerase [Hepatitis B virus]	>...	347	6e-95
gi|23380153|gb|AAM83008.1|	polymerase [Hepatitis B virus]		347	6e-95
gi|23379973|gb|AAM82888.1|	polymerase [Hepatitis B virus]	>...	347	6e-95
gi|23380045|gb|AAM82936.1|	polymerase [Hepatitis B virus]		346	7e-95
gi|23379877|gb|AAM82824.1|	polymerase [Hepatitis B virus]	>...	346	8e-95
gi|23380138|gb|AAM82998.1|	polymerase [Hepatitis B virus]		346	9e-95
gi|23379871|gb|AAM82820.1|	polymerase [Hepatitis B virus]		346	9e-95
gi|23380162|gb|AAM83014.1|	polymerase [Hepatitis B virus]		346	9e-95
gi|23379946|gb|AAM82870.1|	polymerase [Hepatitis B virus]	>...	346	9e-95
gi|23379895|gb|AAM82836.1|	polymerase [Hepatitis B virus]		346	1e-94
gi|23379913|gb|AAM82848.1|	polymerase [Hepatitis B virus]		345	1e-94
gi|23379916|gb|AAM82850.1|	polymerase [Hepatitis B virus]		345	1e-94
gi|23379991|gb|AAM82900.1|	polymerase [Hepatitis B virus]		345	1e-94
gi|23380012|gb|AAM82914.1|	polymerase [Hepatitis B virus]		345	1e-94
gi|23379889|gb|AAM82832.1|	polymerase [Hepatitis B virus]		345	1e-94
gi|23379949|gb|AAM82872.1|	polymerase [Hepatitis B virus]		345	1e-94
gi|23380039|gb|AAM82932.1|	polymerase [Hepatitis B virus]		345	1e-94
gi|23379898|gb|AAM82838.1|	polymerase [Hepatitis B virus]		345	2e-94
gi|23379880|gb|AAM82826.1|	polymerase [Hepatitis B virus]		345	2e-94
gi|23380033|gb|AAM82928.1|	polymerase [Hepatitis B virus]		345	2e-94
gi|23379874|gb|AAM82822.1|	polymerase [Hepatitis B virus]		344	3e-94
gi|23379979|gb|AAM82892.1|	polymerase [Hepatitis B virus]		344	3e-94
gi|23380015|gb|AAM82916.1|	polymerase [Hepatitis B virus]		344	4e-94
gi|23379937|gb|AAM82864.1|	polymerase [Hepatitis B virus]		344	4e-94

```
gi|23379940|gb|AAM82866.1|  polymerase [Hepatitis B virus] >...   343  5e-94
gi|23380054|gb|AAM82942.1|  polymerase [Hepatitis B virus]        343  7e-94
gi|23379910|gb|AAM82846.1|  polymerase [Hepatitis B virus]        343  7e-94
gi|23379901|gb|AAM82840.1|  polymerase [Hepatitis B virus]        343  8e-94
gi|23380018|gb|AAM82918.1|  polymerase [Hepatitis B virus]        343  9e-94
gi|23380027|gb|AAM82924.1|  polymerase [Hepatitis B virus]        342  1e-93
gi|1914714|emb|CAA66689.1|  polymerase [Hepatitis B virus]        342  1e-93
gi|23379982|gb|AAM82894.1|  polymerase [Hepatitis B virus]        342  2e-93
gi|5019986|gb|AAD37962.1|  P protein [Hepatitis B virus]          341  2e-93
gi|23380051|gb|AAM82940.1|  polymerase [Hepatitis B virus]        341  2e-93
gi|27450186|gb|AAO14550.1|AF460223_1  polymerase [Hepatitis ...   341  3e-93
gi|5019949|gb|AAD37932.1|  P protein [Hepatitis B virus]          341  4e-93
gi|27450210|gb|AAO14562.1|AF460235_1  polymerase [Hepatitis ...   338  2e-92
gi|27450206|gb|AAO14560.1|AF460233_1  polymerase [Hepatitis ...   337  4e-92
gi|1107593|emb|CAA56885.1|  polymerase [Hepatitis B virus]        336  7e-92
gi|27450208|gb|AAO14561.1|AF460234_1  polymerase [Hepatitis ...   336  8e-92
gi|27450182|gb|AAO14548.1|AF460221_1  polymerase [Hepatitis ...   336  8e-92
gi|27450184|gb|AAO14549.1|AF460222_1  polymerase [Hepatitis ...   335  2e-91
gi|3820918|emb|CAA08937.1|  polymerase [Hepatitis B virus] >...   332  2e-90
gi|3820942|emb|CAA08951.1|  polymerase [Hepatitis B virus]        330  5e-90
gi|3820933|emb|CAA08947.1|  polymerase [Hepatitis B virus]        326  8e-89
gi|3820945|emb|CAA08953.1|  polymerase [Hepatitis B virus]        326  9e-89
gi|3820930|emb|CAA08945.1|  polymerase [Hepatitis B virus]        325  3e-88

Query= gi|93080|pir||S20757 DNA-directed DNA polymerase (EC
2.7.7.7) - hepatitis B virus (subtype ayw, patient E)
         (832 letters)
cutoff = 3e-88
Database: All non-redundant GenBank CDS
translations+PDB+SwissProt+PIR+PRF
           1,376,942 sequences; 442,405,847 total letters
```

BLASTP 2.2.5 (Nov-16-2002) (Altschul, S.F., et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402 (1997)) against HBV subtype sequence S20757, cutoff = 3e-88 (to select human sequences).

Table 21: CLUSTALW alignment of 19 HBV polymerase sequences

```
                        10        20        30        40        50        60
                        |         |         |         |         |         |
JDVLKS          MPLSYQHFRKLLLLDDGTEAGPLEEELPRLADADLNRRVAEDLNLGNLNVSIPWTHKVGN
JDVLVD          MPLSYQHFRKLLLLDDGTEAGPLEEELPRLADADLHRRVAEDLNLGNLNVSIPWTHKVGN
JDVLJ3          MPLSYQHFRKLLLLDD--EAGPLEEELPRLADEGLNRRVAEDLNLGNLNVSIPWTHKVGN
S71785          MPLSYQHFRKLLLLDD--EAGPLEEELPRLADEDLNRRVAEDLNLGNLNVSIPWTHKVGN
S47406          MPLSYQHFRRLLLLDD--EAGPLEEELPRLADEDLNRRVAEDLNLGNLNVSIPWTHKVGN
JDVLA1          MPLSYQHFRRLLLLDD--EAGPLEEELPRLADEGLNRHVAEELNLGNLNVSIPWTHKVGN
S20752          MPLSYQHFRRLLLLDD--EAGPLEEELPRLADEGLNRRVAEDLNLGNLNVSIPWTHKVGN
S67505          MPLSCPHFRKLLLLDE--EAGPLEEELPRLADEGLNRRVAEDLNLGNLNVSIPWTHKVGN
JQ2229          MPLSYPHFRKLLLLDD--EAGPLEEELPRLADEDLNRRVAADLNLQLPNVSIPWTHKVGN
UNK_494016      ------------------------------------------------------------
UNK_494017      ------------------------------------------------------------
UNK_494015      MPLSYQHFRKLLLLDD--EAGPLEEELPRLADEGLNRRVAEDLNLGNLNVSIPWTHKVGN
UNK_494018      MPLSYQHFRKLLLLDD--EAGPLEEELPRLADEGLNRRVAEDLNLGNLNVSIPWTHKVGN
JDVLVS          MPLSYQHFRKLLLLDD--EAGPLEEELPRLADEDLNRRVAEDLNLGNLNVSIPWTHKVGN
JDVLVR          MPLSYQHFRKLLLLDD--EAGPLEEELPRLADEGLNRRVAEDLNLGNLNVSIPWTHKVGN
T13468          MPLSYQHFRKLLLLDD--EAGPLEEELPRLADEGLNRRVAEDLNLGNLNVSIPWTHKVGN
T13473          MPLSYQHFRKLLLLDD--EAGPLEEELPRLADEDLNRRVAEDLNLGNLNVSIPWTHKVGN
S43491          MPLSYQHFRKLLLLDN--EAGPLEEELPRLADEDLNRRVAEDLNLGNLNVNIPWTHKVGN
S35527          MPLSYQHFRKLLLLDD--EAGPLEEELPRLADEGLNRRVAEDLNLGNLNVSIPWTHKVGN

Prim.cons.      MPLSYQHFRKLLLLDDGTEAGPLEEELPRLADEGLNRRVAEDLNLGNLNVSIPWTHKVGN

                        70        80        90       100       110       120
                        |         |         |         |         |         |
JDVLKS          FTGLYSSTAPIFNPEWQTPSFPKIHLQEDIINRCQQFVGPLTVNEKRRLKLIMPARFYPT
JDVLVD          FTGLYSSTVPIFNPEWQTPSFPKIHLQEDIINRCQQFVGPLTVNEKRRLKLIMPARFYPT
JDVLJ3          FTGLYSSTVPSFNPQWQTPSFPDIHLQEDIINKCKQFVGPLTVNEKRRLKLIMPARFYPN
S71785          FTGLYSSTVPVFNPHWKTPSFPNIHLHQDIIKKCEQFVGPLTVNEKRRLQLIMPARFYPN
S47406          FTGLYSSTVPVFNPHWKTPSFPNIHLRQDIIKKCEQFVGPLTVNEKRRLQLIMPARFYPK
JDVLA1          FTGLYSSTVPVFNPHWKTPSFPNIHLHQDIIKKCEQFVGPLTVNEKRRLQLIMPARFYPK
S20752          FTGFYSSTVPVFNPHWETPSFPNIHLHQDIIKKCEQFVGPLTVNEKRRLQLIMPARFYPK
S67505          FTGLYSSTVPVFNPKWQTPSFPDIHLHQDIINKCEQFVGPLTVNEKRRLKLIIAARFYPN
JQ2229          FTGLYSSTVPAFNPNWSTPSFPDIHLHQDLISKCEQFVGPLTKNELRRLKLVMPARFYPK
UNK_494016      ------------------------------------------------------------
UNK_494017      ------------------------------------------------------------
UNK_494015      FTGLYSSTVPVFNPDWKTPSFPHIHLQEDIINRCQQYVGPLTVNEKRRLKLIMPARFYPN
UNK_494018      FTGLYSSTVPVFNPDWKTPSFPHIHLQEDIINRCQQYVGPLTVNEKRRLKLIMPARFYPN
JDVLVS          FTGLYSSTVPVFNPEWQTPSFPHIHLQEDIINRCQQYVGPLTVNEKRRLKLIMPARFYPK
JDVLVR          FTGLYSSTVPVFNPEWQTPSFPHIHLQEDIINRCQQYVGPLTVNEKRRLKLIMPARFYPN
T13468          FTGLYSSTVPVFNPEWQTPSFPNIHLQEDIINRCQQYVGPLTVNEKRRLKLIMPARFYPN
T13473          FTGLYSSTVPVFNPEWQTPSFPHIHLQEDIINRCQQYVGPLTVNEKRRLKLIMPARFYPN
S43491          FTGLYSSTVPVFNPEWKTPSFPNIHLQEDIIDRCQQYVGPLTVNEKRRLKLIMPARFYPN
S35527          FTGLYSSTVPVFNPECQTPSFPNIHLQEDIINRCQQYVGPLTVNEKRRLKLIMPARFYPN

Prim.cons.      FTGLYSSTVPVFNPEWQTPSFPNIHLQEDIINRCQQFVGPLTVNEKRRLKLIMPARFYPN

                       130       140       150       160       170       180
                        |         |         |         |         |         |
JDVLKS          HTKYLPLDKGIKPYYPDQVVNHYFQTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
JDVLVD          HTKYLPLDKGIKPYYPDQVVNHYFQTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
JDVLJ3          VTKYLPLDKGIKPYYPEHVVNHYFQTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
S71785          VTKYLPLDKGIKPYYPEHLVNHYFQTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
S47406          VTKYLPLDKGIKPYYPEHLVNHYFQTRHYLHTLWKAGILYKRETTHSASFCGSPYSWEQE
JDVLA1          VTKYLPLDKGIKPYYPEHLVNHYFQTRHYLHTLWKAGVLYKRETTHSASFCGSPYSWEQE
S20752          VTKYLPLDKGIKPYYPEHLVNHYFQTRHYLHTLWKAGILYKRETTHSASFCGSPYSWEQD
S67505          APKYLPLDKGIKPYYPEHVVNHYFQTRHYLHILWKAGILYKRETTRSASFCGSPYSWEQE
JQ2229          VTKYFPMDKGIKPYYPEHAVNHYFKTRHYLHTLWKAGILYKRESTRSASFCGSPYSWEQE
```

```
UNK_494016  ------------------------------------------------------------
UNK_494017  ------------------------------------------------------------
UNK_494015  LTKYLPLDKGIKPYYPEYAVNHYFKTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
UNK_494018  LTKYLPLDKGIKPYYPEYAVNHYFKTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
JDVLVS       LTKYLPLDKGIKPYYPEHAVNHYFKTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
JDVLVR       LTKYLPLDKGIKPYYPEHAVNHYFKTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
T13468       LTKYLPLDKGIKPYYPEHAVNHYFKTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
T13473       LTKYLPLDKGIKPYYPEHAVNHYFKTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
S43491       LTKYLPLDKGIKPYYPEHAVNHYFQTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE
S35527       LTKYLPLDKGIKPYYPEHAVNHYFKTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE

Prim.cons.  LTKYLPLDKGIKPYYPEHAVNHYFQTRHYLHTLWKAGILYKRETTRSASFCGSPYSWEQE

                      190         200         210         220         230         240
                       |           |           |           |           |           |
JDVLKS       LQHGRLVIKTSQRHGDESFCSQPSGILSRSSVGPCIRSQLKQSRLGLQPHQGPLASSQPG
JDVLVD       LQHGRLVIKTSQRHGDESFCSQSSGILSRSSVGPCIRSQLKQSRLGLQPRQGRLASSQPS
JDVLJ3       LQHGRLVLQTSTRHGDKSFRPQSSGILSRSPVGPCIQSQLRQSRLGPQPTQGQLAGLQQG
S71785       LQHG-----------AESFHQQSSGILSRPPVGSSLQSKHRKSRLGLQSQQGHLARRQQG
S47406       LQHG-----------AESFHQQSSGILSRPPVGSSLQSKHSKSRLGLQSQQGHLARRQQG
JDVLA1       LQHG-----------AESFHQQSSGILSRPPVGSSLQSKHCKSRLGLQSQQGLLARRQQG
S20752       LQHG-----------AESIHQQSSGILSRPPVGSSLQSKHRKSRLGLQSQQGHLARRQQG
S67505       LQHG-----------AEPVCQQSLGILPRASVGSPVQSQLKQSRLGLQSQQRQLARSHQG
JQ2229       LQHGSTSLNDTKRHGTESLCAQSSGILSRPSAGSAIQSKFQQSRLGLQHKQGQLANGKQG
UNK_494016  ------------------------------------------------------------
UNK_494017  ------------------------------------------------------------
UNK_494015  LQHGRLVFQTSTRHGDESFCSQSSGILSRSPVGPCVRSQLKQSRLGLQPQQGSLARGKSG
UNK_494018  LQHGRLVFQTSTRHGDESFCSQSSGILSRSPVGPCVRSQLKQSRLGLQPQQGSLARGKSG
JDVLVS       LQHGRLVFQTSTRHGDESFCSQSSGILSRSPVGPCVRSQLKQSRLGLQPQQGSLARGKSG
JDVLVR       LQHGRLVFQTSTRHGDESFCSQSSGILSRSPVGPCVRSQLKQSRLGLQPQQGSMARGKSG
T13468       LQHGRLVFQTSTRHGDESFCSQSSGILSRSPVGPGVRSQLKQSRLGLQPQQGSLARGKSG
T13473       LQHGRLVFQTSTRHGDKSFCSQSSGILSRSPVGPCVRSQLKQSRLGLQPQQGSLARGQSG
S43491       LQHGRLVFQTSTRHGDESFCSQSSGILSRSPVGPCVRSQLKQSRLGLQPQQGSLARGKAG
S35527       LQHGRLVFQTSTRHGDESFCSQSSGILSRSPVGPCVRSQLKQSRLGLQPQQGSLARRNQG

Prim.cons.  LQHGRLVFQTSTRHGDESFCSQSSGILSRSPVGPCVRSQLKQSRLGLQPQQGSLARGQQG

                      250         260         270         280         290         300
                       |           |           |           |           |           |
JDVLKS       RSGSIRARVHPSTRRCFGVEPSGSGHVDPSVNNSSSCLRQSAVRKAAYSHLSTSKRQSSS
JDVLVD       RSGSIRAKAHPSTRRYFGVEPSGSGHIDHSVNNSSSCLHQSAVRKAAYSHLSTSKRQSSS
JDVLJ3       GSGSIRAGIHSTPWGTVGVEPSSSGHTHNCANSSSSCLHQSAVRKEAYSPVSTSKRHSSS
S71785       RSWSIRAGIHPTARRPFGVEPSGSGHNTNLASKSASCIYQSPVRKAAYPAVSTFEKHSSS
S47406       RSWSIRAGIHPTARRPFGVEPSGSGHNTNLASKSASCLYQSPVRKAAYPAVSTFEKHSSS
JDVLA1       RSWSIRAGIHPTARRPFGVEPSGSGHTTNLASKSASCLHQSPVRKATYPSVSTFEKHSSS
S20752       WSWSIRAGTHPTARRPFGVEPSGSGHTTHRASKSASCLYQSPDRKATYPSVSTFERHSSS
S67505       RSGSIRARVHSTTRRSFRVELSGSGSNHNIASTSSSCRHQSAVRETAYSHLSTVERHSSS
JQ2229       RSGRLRSRVHTPTRWPAGVEPSSTRCVNNLASRSASCFHQSAVREKANPSLSTSKRHTST
UNK_494016  ------------------------------------------------------------
UNK_494017  ------------------------------------------------------------
UNK_494015  RSGSIWSRVHPTTRRPFGVEPSGSGHIDNTASSTSSCLHQSAVRKTAYSHLSTSKRQSSS
UNK_494018  RSGSIWSRVHPTTRRPFGVEPSGSGHIDNTASSTSSCLHQSAVRKTAYSHLSTSKRQSSS
JDVLVS       RSGSIRARVPPTTRRSFGVEPSGSGHIDNRASSTSSCLHQSAVRKTAYSHLSTSKRQSSS
JDVLVR       RSGSIRARVHPTTRRSFGVEPSGSGHIDNSASSTSSCLHQSAVRKTAYSHLSTSKRQSSS
T13468       RSGSIRARVHPTTRRSFGVEPSGSGHIDNSARSASSCLHQSAVRKTAYSHLSTSKRQSSS
T13473       RIGSIRARVHPTTRRSFGVEPSGSGHIDNSASSPSSCLHQSAVRKTAYCHLSTTKRQSSS
S43491       RSGSIRARVHPTTRRPFGVEPSGSGHIDNSASSASSCFHQSAVRKTAYSHLSTSKRQSSS
S35527       RSGRLRARVHPTTRRSFGVEPSGSGHLDNSASSSSSCLHQSAVRKTAYSHLSTSKRQSSS

Prim.cons.  RSGSIRARVHPTTRRPFGVEPSGSGHIDNSASSSSSCLHQSAVRKTAYSHLSTSKRQSSS
```

```
                        310       320       330       340       350       360
                        |         |         |         |         |         |
JDVLKS          GHAVEFHCLPPSSARPQSQGSVFSCWWLQFRNSKPCSEYCLSHLVNLREDRGPCDEHGEH
JDVLVD          GHAVEFHCLPPNSAGSQSQGSVSSCWWLQFRNSKPCSEYCLSHLVNLREDRGPCDEHGEH
JDVLJ3          GNAVELHHVPPNSSRSQSQGSVLSCWWLQFRNSKPCSEHCLFHIVNLIDDWGPCAEHGEH
S71785          GHAVELHNFPPNSARSQGERPVFPCWWLQFRNSKPCSDYCLSHIVNLLEDWGPCTEHGEH
S47406          GHAVELHNLPPNSSRSQGERPVFPCWWLQFRNSKPCSDYCLSHIVNLLEDWGPCTEHGEH
JDVLA1          GHAVELHNLPPNSARSQSERPVSPCWWLQFRNSKPCSDYCLSHIVNLLEDWGPCAEHGEH
S20752          GRAVELHNFPPNSARSQSERPIFPCWWLQFRNSKPCSDYCLSHIVNLLEDWGPCAEHGEH
S67505          GHEVELYSIPPNSARSQSTGPILSCWWLQFRNSEPCSDYCLSHLVNLLEDWGPCDEYGEH
JQ2229          GNAVELNPVPPSSVGSQGKGSVLPCWWLQFRDTEPCSDYCLSHIINLLEDWGPCYEHGQH
UNK_494016      -----LHNIPPSSARSQSEGPIFSCWWLQFRNSKPCSDYCLTHIVNLLEDWGPCTEHGEH
UNK_494017      -----LHNIPPSSARSQSEGPIFSCWWLQFRNSKPCSDYCLTHIVNLLEDWGPCTEHGEH
UNK_494015      GHAVE----------------------------------------------------------
UNK_494018      GHAVE----------------------------------------------------------
JDVLVS          GHAVELHHISPSPARSQSEGPIFSSWWLQFRNSKPCSDYCLTHIVNLLEDWGPCTEHGEH
JDVLVR          GHAVEFHNIPPSSARSQSEGPIFSCWWLQFRNSKPCSDYCLTHIVNLLEDWGPCTEHGEH
T13468          GHAVELHPCWWLQFRNSKPCSDYCLTHIVNLLEDWGPCTEHGEHNIRIPRTPARVTGGVF
T13473          GHAVETCWWLQFRNSKPCSDYCLTHIVNLLEDWGPCTEHGEHNIRIPRTPARVTGGVFLV
S43491          GHAVELHNIPPSSARSKSEGPLFPCWWLQFRNSKPCSDYCLTHIVNLLEDWGPCTEHGEH
S35527          GHAVELHNIPPSSARSQSEGPIFSCWWLQFRNSKPCSDYCLTHIVNLLEDWGPCTEHGEH

Prim.cons.      GHAVELHNIPPSSARSQSEGP2FSCWWLQFRNSKPCSDYCLSHIVNLLEDWGPCTEHGEH

                        370       380       390       400       410       420
                        |         |         |         |         |         |
JDVLKS          HIRIPRTPARVTGGVFLVDKNPHNTAESRLVVDFSQFSRGITRVSWPKFAIPNLQSLTNL
JDVLVD          HIRIPRTPARVTGGVFLVDKNPHNTAESRLVVDFSQFSRGISRVSWPKFAVPNLQSLTNL
JDVLJ3          RIRTPRTPARVTGGVFLVDKNPHNTSESRLVVDFSQFSRGNTRVSWPKFAVPNLQSLTNL
S71785          HIRIPRTPARVTGGVFLVDKNPHNTAESRLVVDFSQFSRGNHRVSWPKFAVPNLQSLTNL
S47406          HIRIPRTPARVTGGVFLVDKNPHNTAESRLVVDFSQFSRGNYRVSWPKFAVPNLQSLTNL
JDVLA1          HIRIPRTPARVTGGVFLVDKNPHNTEESRLVVDFSQFSRGNHRVSWPKFAVPNLQSLTNL
S20752          HIRIPRTPARVTGGVFLVDKNPHNTAESRLVVDFSQFSRGNYRVSWPKFAVPNLQSLTNL
S67505          HIRIPRTPARVTGGVFLVDKNPHNTTESRLVVDFSQFSRGSTRVSWPKFAVPNLQSLTNL
JQ2229          YIRTPRTPARVTGGVFLVDKNPHNTTESRLVVDFSQFSRGTTRVSWPKFAVPNLQSLTNL
UNK_494016      NIRIPRTPARVTGGVFLVDKNPHNTTESRLVVDFSQFSRGSTHVSWPKFAVPNLQSLTNL
UNK_494017      NIRIPRTPARVTGGVFLVDKNPHNTTESRLVVDFSQFSRGSTHVSWPKFAVPNLQSLTNL
UNK_494015      ----------------------------------------------------------
UNK_494018      ----------------------------------------------------------
JDVLVS          NIRIPRTPARVTGGVFLVDKNPHNTTESRLVVDFSQFSRGSTHVSWPKFAVPNLQSLTNL
JDVLVR          NIRIPRTPARVTGGVFLVDKNPHNTTESRLVVDFSQFSRGSTHVSWPKFAVPNLQSLTNL
T13468          LVDKNPHNTTESRLVVDFSQFSRGSTX----------------VSWPKFAVPNLQSLTNL
T13473          DKNPHNTTESX-----------------LVVDFSQFSRGSTQVSWPKFAVPNLQSLTNL
S43491          NIRIPRTPARVTGGVFLVDKNPHNTTESRLVVDFSQFSRGNTQVSWPKFAVPNLQSLTNL
S35527          NIRIPRTPARVTGGVFLVDKNPHNTTESRLVVDFSQFSRGSTHVSWPKFAVPNLQSLTNL

Prim.cons.      HIRIPRTPARVTGGVFLVDKNPHNTTESRLVVDFSQFSRGSTRVSWPKFAVPNLQSLTNL

                        430       440       450       460       470       480
                        |         |         |         |         |         |
JDVLKS          LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLIGSSGLSRYVARLSSNSRINNNQYGTMQNL
JDVLVD          LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLIGSSGLSRYVARLSSNSRINNNQYGTMQNL
JDVLJ3          LSSDLSWLSLDVSAAFYHLPLHPAAMPHLLVGSSGLSRYVARLSSNSRINNNQYGTMQNL
S71785          LSSNLSWLSLDVSAAFYHLPLHPASMPHLLVGSTGLSRYVARVSSNSRIINHQHRTMQNL
S47406          LSSNLSWLSLDVSAAFYHLPLHPAAMPHLLVGSSGLSRYVARVSSNSRIFNHQRGTMQNL
JDVLA1          LSSNLSWLSLDVSAAFYHLPLHPAAMPHLLVGSSGLSRYVARLSSNSRIFNNQRGTMQNL
S20752          LSSNLSWLSLDVSAGFYHLPLHPAAMPHLLVGSSGVSRYVARLSSNSRNNNNQYGTMQNL
S67505          LSSNLSWLSLDVSAAFYHLPLHPAAMPHLLVGSSGLSRYVARLSSTSRIIDHQHGTMQNL
JQ2229          LSSNLSWLSLDVSAAFYHLPLHPAAMPHLLVGSSGLSRYVARLSSTSRIHDHQHGTLQNL
UNK_494016      LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLVGSSGLPRYVARLSSTSRNINYQHGTMQNL
UNK_494017      LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLVGSSGLPRYVARLSSTSRNINYQHGTMQNL
```

```
UNK_494015  ------------------------------------------------------------
UNK_494018  ------------------------------------------------------------
JDVLVS      LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLVGSSGLPRYVARLSSTSRNINHQHGTMQDL
JDVLVR      LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLVGSSGLPRYVARLSSTSRNINHQHGTMQDL
T13468      LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLVGSSGLPRYVARLSSTSRNINHQHGAMQDL
T13473      LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLVGSSGLPRYVARLSSTSRNXNYQHGTMQDL
S43491      LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLVGSSGLPRYVARLSSTSRNINXQHGTMQDL
S35527      LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLVGSSGLQRYVARLSSTSRNINYQHGTMQDL

Prim.cons.  LSSNLSWLSLDVSAAFYHIPLHPAAMPHLLVGSSGLSRYVARLSSTSRNINHQHGTMQNL

                      490         500         510         520         530         540
                       |           |           |           |           |           |
JDVLKS      HDSCSRQLYVSLMLLYKTYGWKLHLYSHPIVLGFRKIPMGVGLSPFLLAQFTSAICSVVR
JDVLVD      HDSCSRQLYVSLMLLYKTYGWKLHLYSHPIVLGFRKIPMGVGLSPFLLAQFTSAICSVVR
JDVLJ3      HDSCSRNLYVSLMLLYKTYGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR
S71785      HDYCSRNLYVSLLLLYQTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR
S47406      HDYCSRNLYVSLLLLYQTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR
JDVLA1      HDSCSRNLYVSLLLLYQTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR
S20752      HDSCSRQLYVSLMLLYQNFGWKLHLYSHPIVLGFRKIPMGVGLSPFLLAQFTSAICSVVR
S67505      HDHCSRNLFVSLMLLYKTFGRKLHLYSHPIVLGFRKIPMGVGLSPFLLAQFTSAICSVVR
JQ2229      HNSCTRNLYVSLLLLFQTLGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSSICSVVR
UNK_494016  HDSCSRNLYVSLLLLYKTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR
UNK_494017  HDSCSRNLYVSLLLLYKTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR
UNK_494015  ------------------------------------------------------------
UNK_494018  ------------------------------------------------------------
JDVLVS      HDSCSRNLYVSLLLLYKTFGRKLHLYSHPIILGFRKIPMGGGLSPFLLAQFTSAICSVVR
JDVLVR      HDSCSRNLYVSLLLLYKTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR
T13468      HESCSRNLYVSLLLLYKTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR
T13473      HDSCSKHLYVSLLLLYKTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR
S43491      HDSCSKHLYVSLLLLYKTFGRKLHLYSHPIXLGFRKIPMGVGLSPFLLAQFTSAICSVVR
S35527      HDSCSRNLYVSLLLVYKTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVC

Prim.cons.  HDSCSRNLYVSLLLLYKTFGRKLHLYSHPIILGFRKIPMGVGLSPFLLAQFTSAICSVVR

                      550         560         570         580         590         600
                       |           |           |           |           |           |
JDVLKS      RAFPHCLAFSYMDDVVLGAKSVQHREFLYTAVTNFLLSLGIHLNPNKTKRWGYSLNFMGY
JDVLVD      RAFPHCLAFSYMDDVVLGAKSVQHRESLYTAVTNFLLSLGIHLNPNKTKRWGYSLNFMGY
JDVLJ3      RAFPHCLAFSYMDDVVLGAKSVQHRESLYTAVTNFLLSLGIHLNPNKTKRWGYSLNFMGY
S71785      RAFPHCLAFSYMDDVVLGAKSVQHLESLYAAVTNFLLSLGIHLNPQKTKRWGYSLNFMGY
S47406      RAFPHCLAFSYMDDVVLGAKSVQHLESLFTAVTNFLLSLGIHLNPNKTKRWGYSLHFMGY
JDVLA1      RAFPHCLAFSYMDDVVLGAKSVQHLESLFTAVTNFLLSLGIHLNPNKTKRWGYSLNFMGY
S20752      RAFPHCLAFSYMDDVVLGAKTVHHLESLFTAVTNFLLSLGIHLNPNKTKRWGYSLHFMGY
S67505      RAFPHCLAFSYMDDVVLGAKSVQHLESLFTAVTNFLLSLGIHLNPNKTKRWGYSLHFMGY
JQ2229      RAFPHCLAFSYMDDLVLGAKSVQHLESIYTAVTNFLLSLGIHLNPNKTKRWGYSLNFMGY
UNK_494016  RAFPHCLAFSYMDDLVLGAKSVQHLESLYTAVTNFLLSVGIHLNTSKTKRWGYSLHFMGY
UNK_494017  RAFPHCLAFSYMDDVVLGAKSVQHLESLFTSITNFLLSLGIHLNPNKTKRWGYSLNFMGY
UNK_494015  ------------------------------------------------------------
UNK_494018  ------------------------------------------------------------
JDVLVS      RAFPHCLAFSYMDDVVLGAKSVQHLESLFTSITNFLLSLGIHLNPNKTKRWGYSLNFMGY
JDVLVR      RAFPHCLAFSYMDDVVLGAKSVQHLESLFTSITNFLLSLGIHLNPNKTKRWGYSLNFMGY
T13468      RAFPHCLAFSYMDDVVLGAKSVQHLESLFTAITNFLLSLGIHLNPHKTKRWGYSLNFMGY
T13473      RAFPHCLAFSYMDDVVLGAKSVQHLEALYTSITNFLLSLGIHLNPNKTKRWGYSLNFMGY
S43491      RAFPHCLAFSYMDDVVLGAKSVQHLESLFTSITNFMLSLGIHLNPNKTKRWGYSLNFMGY
S35527      RAFPHCLAFSYMDDVVLGAKSVQHLESLFTSITNFLLSLGIHLNPNKTKRWGYSLNFMGY

Prim.cons.  RAFPHCLAFSYMDDVVLGAKSVQHLESLFTAVTNFLLSLGIHLNPNKTKRWGYSLNFMGY

                      610         620         630         640         650         660
                       |           |           |           |           |           |
```

```
JDVLKS        VIGSWGTLPQDHIVQKIKHCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
JDVLVD        IIGSWGTLPQDHIVQKIKHCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
JDVLJ3        VIGSWGTLPQEHIVLKIKQCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
S71785        VIGSYGSLPQDHIIQKIKECFRKLPINRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
S47406        VIGCYGSLPQDHIIQKIKECFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
JDVLA1        VIGCYGSLPQDHIIQKIKECFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALKPLY
S20752        VIGCYGSLPQEHIIQKIKECFRKVPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
S67505        IIGSWGSLPQDHIVQKIKQCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
JQ2229        VIGSWGSLPQDHIVHKIKECFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
UNK_494016    VIGSWGTLPQEHIVQKLKQCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
UNK_494017    VIGSWGTLPQEHIVQKLKQCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
UNK_494015    ------------------------------------------------------------
UNK_494018    ------------------------------------------------------------
JDVLVS        VIGSWGTLPQEHIVLKIKQCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
JDVLVR        VIGSWGTLPQEHIVLKLKQCFRKLPVNSPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
T13468        VIGSWGTLPQEHIVQKIKQCFRKLPVNRPIDWKVCQXIVGLLGFAAPFTQCGYPALMPLY
T13473        VIGSWGTLPQDHIVQKLKQCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
S43491        VIGSWGTLPQEHIVQKLKQCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY
S35527        VIGSWGTLPQEHIVLKLKQCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY

Prim.cons.    VIGSWGTLPQEHIVQKIKQCFRKLPVNRPIDWKVCQRIVGLLGFAAPFTQCGYPALMPLY

                        670       680       690       700       710       720
                         |         |         |         |         |         |
JDVLKS        ACIQAKQAFTFSPTYKAFLSKQYMNLYPVARQRPGLCQVFADATPTGWGLAIGHQRMRGT
JDVLVD        ACIQAKQAFTFSPTYKAFLSKQYMNLYPVARQRPGLCQVFADATPTGWGLAIGHQRMRGT
JDVLJ3        ACIQAKQAFTFSPTYKAFLNKQYLNLYPVARQRPGLCQVFADATPTGWGLAIGHQRMRGT
S71785        ACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRPGLCQVFADATPTGWGLVMGHQRMRGT
S47406        ACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRPGLCQVFADATPTGWGLVMGHQRMRGT
JDVLA1        ACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRPGLCQVFADATPTGWGLVMGHQRMRGT
S20752        ACIQFKQAFTFSPTYKAFLCKQYLNLYPVARQRPGLCQVFADATPTGWGLGMGHQRMRGT
S67505        ACIQAKQAFTFSPTYKAFLRTQYLTLYPVARQRPGLCQVFADATPTGWGLAIGHQRMRGT
JQ2229        ACITAKQAFVFSPTYKAFLCKQYMNLYPVARQRPGLCQVFADATPTGWGLAIGHQRMRGT
UNK_494016    ACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRSGLCQVFADATPTGWGLAIGHRRMRGT
UNK_494017    ACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRSGLCQVFADATPTGWGLAIGHRRMRGT
UNK_494015    ------------------------------------------------------------
UNK_494018    ------------------------------------------------------------
JDVLVS        ACIQSKQAFTFSPTYKAFLCKQYLHLYPVARRTALCQVFADATPTGWGLAIGHRRMRGTF
JDVLVR        ACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRSGLCQVFADATPTGWGLAIGHRRMRGT
T13468        ACIQAKQAFTFSPTYKAFLCKQYLHLYPVARQRSGLCQVFADATPTGWGLAIGHRRMRGT
T13473        ACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRSGLCQVFADATPTGWGLAIGHRRMRGT
S43491        ACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRSGLCQVFADATPTGWGLAIGHQRMRGT
S35527        ACIQSKQAFTFSPTYKAFLCQQYLHLYPVARQRSGLCQVFGDATPTGWGLAIGHRRMRGT

Prim.cons.    ACIQSKQAFTFSPTYKAFLCKQYLNLYPVARQRPGLCQVFADATPTGWGLAIGHQRMRGT

                        730       740       750       760       770       780
                         |         |         |         |         |         |
JDVLKS        FVAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
JDVLVD        FVAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCTANWILRGTSFV
JDVLJ3        FVSPLPIHTVELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
S71785        FLAPLPIHTAELLAACFARSRSGANILGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
S47406        FLAPLPIHTAELLAACFARSRSGANILGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
JDVLA1        FQAPLPIHTAELLAACFARSRSGANILGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
S20752        FSAPLPIHTAELLAACFARSRSGANILGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
S67505        FVAPLPIHTAELLAACFARSRSGANIIGTDNSVVLSPKYTSFPWLLGCAANWILRRTSFV
JQ2229        FVAPLPIHTAELLAACFARSRSGATLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
UNK_494016    FVAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
UNK_494017    FVAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
UNK_494015    ------------------------------------------------------------
UNK_494018    ------------------------------------------------------------
```

```
JDVLVS        VAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTYFVY
JDVLVR        FVAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
T13468        FVAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
T13473        FVAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
S43491        FVAPLPIHTAELLAACFARSRSGATLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV
S35527        FVAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV

Prim.cons.    FVAPLPIHTAELLAACFARSRSGAKLIGTDNSVVLSRKYTSFPWLLGCAANWILRGTSFV

                       790         800         810         820         830         840
                        |           |           |           |           |           |
JDVLKS        YVPSALNPADDPSRGRLGLSRPLLRLPFQPTTGRTSLYAVSPSVPSHLPVRVHFASPLHV
JDVLVD        YVPSALNPADDPSRGRLGLSRPLLRLPFQPTTGRTSLYAVSPSVPSHLPVRVHFASPLHV
JDVLJ3        YVPSALNPADDPSRGRLGLYRPLLRLPYRPTTGRTSLYADSPSVPSHLPDRVHFASPLHV
S71785        YVPSALNPADDPSRGRLGIFRPLLRLPFRPTTGRTSLYADSPSVPSHLPVRVHFASPLHV
S47406        YVPSALNPADDPSRGRLGLSRPLLRLPFRPTTGRTSLYADSPSVPSHLPDRVHFASPLHV
JDVLA1        YVPSALNPADDPSRGRLGLSRPLLRLPFRPTTGRTSLYADSPSVPSHLPDRVHFASPLHV
S20752        YVPSALNPADDPSRGRLGLSRPLLCLPFRPTTGRTSLYADSPSVPSHLPDRVHFASPLHV
S67505        YVPSALNPADDPSRGRLGLYRPLLRPWFRPTTGRTSLYAVSPSVPSHLPVRVHFASPLHV
JQ2229        YVPSALNPADDPSRGRLGLYRPLLRLPFQPTTGRTSLYADSPSVPSHLPDRVHFASPLHV
UNK_494016    YVPSALNPADDPSRGRLGLYRPLLHLPFRPTTGRTSLYAVSPSVPSHLPDRVHFASPLHV
UNK_494017    YVPSALNPADDPSRGRLGLYRPLLHLPFRPTTGRTSLYAVSPSVPSHLPDRVHFASPLHV
UNK_494015    ------------------------------------------------------------
UNK_494018    ------------------------------------------------------------
JDVLVS        VPSALNPADDPSRGRLGLIRPLLHLRFRPTTGRTSLYAVSPSVPSHLPDRVHFASPLHVA
JDVLVR        YVPSALNPADDPSRGRLGLYRPLLLLPFRPTTGRTSLYAVSPSVPSHLPDRVHFASPLHV
T13468        YVPSALNPADDPSRGRLGLYRPLLHLPFRPTTGRTSLYAVSPSVPSHLPDRVHFASPLHV
T13473        YVPSALNPADDPSRGRLGLYRPLLHLPFRPTTGRTSLYAVSPSVPSHLPDRVHFASPLHV
S43491        YVPSALNPADDPSRGRLGLYRPLLRLSFRPTTGRTSLYAVSPSVPSHLPDRVHFASPLHV
S35527        YVPSALNPADDPSRGRLGLYRPLLHLPFQPTTGRTSLYAVSPSVPSHLPVRVHFASPLHV

Prim.cons.    YVPSALNPADDPSRGRLGLYRPLLRLPFRPTTGRTSLYAVSPSVPSHLPDRVHFASPLHV


JDVLKS         AWRPP
JDVLVD         AWRPP
JDVLJ3         AWRPP
S71785         AWRPP
S47406         AWRPP
JDVLA1         AWRPP
S20752         AWRPP
S67505         AWRPP
JQ2229         AWRPP
UNK_494016     AWRPP
UNK_494017     AWRPP
UNK_494015     -----
UNK_494018     -----
JDVLVS         WRPP-
JDVLVR         AWRPP
T13468         AWRPP
T13473         AWRPP
S43491         AWRPP
S35527         AWRPP

Prim.cons.     AWRPP
```

CLUSTALW alignment of 19 HBV polymerare sequences representing the sybtypes adw (4), ayw (5), ayr (4) and adr (6) (NPS@: Network Protein Sequence Analysis, TIBS Vol. 25, No 3 (291):147-150, Combet C., Blanchet C., Geourjon C. and Deléage G. (March 2000))

```
CLUSTALW options used :
endgaps=1
gapdist=8
gapext=0.2
gapopen=10.0
hgapresidues=GPSNDQERK
ktuple=1
matrix=gonnet
maxdiv=30
outorder=aligned
pairgap=3
score=percent
topdiags=5
type=PROTEIN
window=5
```

Table 22. HCV Multiple Sequence Alignment
GCG Multiple Sequence File.
Written by Omiga 1.1

Name: BEBE1	Len:	3052	Check: 2605	Weight:	1.00
Name: D89815	Len:	3052	Check: 9655	Weight:	1.00
Name: ED43type_4	Len:	3052	Check: 4987	Weight:	1.00
Name: HC_C2	Len:	3052	Check: 6273	Weight:	1.00
Name: HC_G9	Len:	3052	Check: 217	Weight:	1.00
Name: HCU16326	Len:	3052	Check: 5854	Weight:	1.00
Name: HCV_H_CMR	Len:	3052	Check: 4932	Weight:	1.00
Name: HCV_J1	Len:	3052	Check: 4947	Weight:	1.00
Name: HCV_J483	Len:	3052	Check: 2553	Weight:	1.00
Name: HCV_J8	Len:	3052	Check: 9778	Weight:	1.00
Name: HCV_JK1	Len:	3052	Check: 4917	Weight:	1.00
Name: HCV_JS	Len:	3052	Check: 3982	Weight:	1.00
Name: HCV_K1_R1	Len:	3052	Check: 9084	Weight:	1.00
Name: HCV_K1_R2	Len:	3052	Check: 47	Weight:	1.00
Name: HCV_K1_R3	Len:	3052	Check: 1630	Weight:	1.00
Name: HCV_K1_S1	Len:	3052	Check: 3578	Weight:	1.00
Name: HCV_K1_S2	Len:	3052	Check: 9909	Weight:	1.00
Name: HCV_K1_S3	Len:	3052	Check: 9508	Weight:	1.00
Name: HCV_L2	Len:	3052	Check: 4175	Weight:	1.00
Name: HCV_N	Len:	3052	Check: 1702	Weight:	1.00
Name: HCV12083	Len:	3052	Check: 7564	Weight:	1.00
Name: HCV1480	Len:	3052	Check: 5620	Weight:	1.00
Name: HCVPOLYP	Len:	3052	Check: 2663	Weight:	1.00
Name: HD_1	Len:	3052	Check: 4040	Weight:	1.00
Name: HPCCGAA	Len:	3052	Check: 5414	Weight:	1.00
Name: HPCFG	Len:	3052	Check: 7119	Weight:	1.00
Name: HPCGENANTI	Len:	3052	Check: 9591	Weight:	1.00
Name: HPCGENOM	Len:	3052	Check: 2009	Weight:	1.00
Name: HPCHUMR	Len:	3052	Check: 4863	Weight:	1.00
Name: HPCJ	Len:	3052	Check: 3553	Weight:	1.00
Name: HPCJCG	Len:	3052	Check: 6658	Weight:	1.00
Name: HPCJK046	Len:	3052	Check: 436	Weight:	1.00
Name: HPCJK049	Len:	3052	Check: 9796	Weight:	1.00
Name: HPCJTA	Len:	3052	Check: 2902	Weight:	1.00
Name: HPCJTB	Len:	3052	Check: 4237	Weight:	1.00
Name: HPCK3A	Len:	3052	Check: 2180	Weight:	1.00
Name: HPCPLYPRE	Len:	3052	Check: 6557	Weight:	1.00
Name: HPCPOLP	Len:	3052	Check: 1218	Weight:	1.00
Name: HPCPP	Len:	3052	Check: 3845	Weight:	1.00
Name: HPCUNKCD	Len:	3052	Check: 6214	Weight:	1.00
Name: MKC1A	Len:	3052	Check: 1615	Weight:	1.00
Name: NDM59	Len:	3052	Check: 9717	Weight:	1.00
Name: NZLI	Len:	3052	Check: 2255	Weight:	1.00
Name: SA13	Len:	3052	Check: 9158	Weight:	1.00
Name: Th580	Len:	3052	Check: 37	Weight:	1.00
Name: Type_3a_CB	Len:	3052	Check: 7958	Weight:	1.00
Name: TypeV_D	Len:	3052	Check: 303	Weight:	1.00
Name: VN004	Len:	3052	Check: 3873	Weight:	1.00
Name: VN235	Len:	3052	Check: 7756	Weight:	1.00
Name: VN405	Len:	3052	Check: 7249	Weight:	1.00

```
            1                                                   50
    BEBE1   MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRAAR
    D89815  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
ED43type_4  MSTNPKPQRK TKRNTNRRPM DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
    HC_C2   MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
    HC_G9   MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRVGVRATR
```

```
HCU16326   MSTNPKPQRK TKRNTNRRPQ DIKFPGGGQI VGGVYLLPRR GPRLGVRATR
HCV_H_CMR  MSTNPKPQRK TKRNTNRRPQ DVEFPGGGQI VGGVYLLPRR GPRLGVRATR
   HCV_J1  MSTIPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
 HCV_J483  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
   HCV_J8  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
  HCV_JK1  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
   HCV_JS  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
HCV_K1_R1  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
HCV_K1_R2  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
HCV_K1_R3  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
HCV_K1_S1  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
HCV_K1_S2  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
HCV_K1_S3  MSTNPKPQRQ TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
   HCV_L2  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
    HCV_N  MSTNPKPQRK TKRNTNRRPQ EVKFPGGGQI VGGVYLLPRR GPRLGVRATR
 HCV12083  MSTLPKPQRK TKRNTNRRPM DVKFPGGGQI VGGVYLLPRR GPRLGVRAIR
  HCV1480  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRK GPRLGVRATR
 HCVPOLYP  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPKLGVRATR
     HD_1  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRALR
  HPCCGAA  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
    HPCFG  MSTLPKPKRQ TKRNTLRRPK NVKFPAGGQI VGEVYVLPRR GPQLGVREVR
HPCGENANTI MSTNGKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
  HPCGENOM  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
  HPCHUMR  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRAPR
     HPCJ  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
   HPCJCG  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
 HPCJK046  MSTNPKPQRQ TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
 HPCJK049  MSTLPKPQRI TKRNINRRPQ DVKFPGGGQI VGGVYVLPRR GPKLGVRAVR
   HPCJTA  MSTNPKPQRK TKRNTYRRPQ DVKFPGGGQI VGGVYVLPRR GPTLGVRATR
   HPCJTB  MSTNPKPQRK TKRNTYRRPQ DVKFPGGGQI VGGVYVLPRR GPTLGVRATR
   HPCK3A  MSTLPKPQRK TKRNTIRRPQ DVKFPGGGVI YVGVYVLPRR GPRLGVRATR
 HPCPLYPRE MSTNPKPQKK NKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
  HPCPOLP  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
    HPCPP  MSTNPKPQRK IKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
 HPCUNKCD  MSTNPKPQRK TKRNTNRRPQ DIKFPGGGQI VGGVYLLPRR GPRLGVRATR
    MKC1A  MSTNPKPQRK IKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
    NDM59  MSTNPKPQRK TKRNTSRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
     NZLI  MSTLPKPQRK TKRNTIRRPQ DVKFPGGGQI VGGVYVLPRR GPRLGVRATR
     SA13  MSTNPKPQRK TKRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
     Th580  MSTLPKPQRK TKRNTNRRPM DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
Type_3a_CB MSTLPKPQRK TKRNTIRRPQ DVKFPGGGQI VGGVYVLPRR GPRLGVCATR
  TypeV_D  MSTLPKPQRK TKRNTIRRPQ DVKFPGGGQI VGGVYVLPRR GPRLGVRATR
    VN004  MSTLPKPQRK TKRNTNRRPM DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
    VN235  MSTLPKPQKR NQRNTNRRPQ DVKFPGGGQI VGGVYLLPRR GPRLGVRATR
    VN405  MSTLPKPQRK TKRNTNRRPM DVKFPGGGQI VGGVYLLPRR GPRLGVRATR

           51                                                  100
    BEBE1  KTSERSQPRG RRQPIPKDRR STGKSWGRPG YPWPLYRNEG LGWAGWLLSP
    D89815 KTSERSQPRG RRQPIPKARR PEGRTWAQPG YPWPLYGNEG LGWAGWLLSP
ED43type_4 KTSERSQPRG RRQPIPKARR PEGRSWAQPG YPWPLYGNEG CGWAGWLLSP
     HC_C2 KTSERSQPRG RRQPIPKARR PEGRAWAQPG YPWPLYGNEG MGWAGWLLSP
     HC_G9 KTSERSQPRG RRQPIPKARR PEGRSWAQPG YPWPLYGNEG CGWAGWLLSP
  HCU16326 KTSERSQPRG RRQPIPKARR PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
 HCV_H_CMR KTSERSQPRG RRQPIPKARR PEGRTWAQPG YPWPLYGNEG CGWAGWLLSP
   HCV_J1  KTSERSQPRG RRQPIPKVRR PEGRTWAQPG YPWPLYGNEG CGWAGWLLSP
 HCV_J483  KTSERSQPRG WRQPIPKARR PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
   HCV_J8  KTSERSQPRG RRQPIPKDRR STGKSWGKPG YPWPLYGNEG CGWAGWLLSP
  HCV_JK1  KTSERSQPRG RRQPIPKARQ PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
   HCV_JS  KTSERSQPRG RRQPIPKARR PEGRTWAQPG YPWPLYGNEG MGWAGWLLSP
 HCV_K1_R1 KTSERSQPRG RRQPIPKARR PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
 HCV_K1_R2 KTSERSQPRG RRQPIPKARQ PEGRAWAQPG YPWPLYGNEG MGWAGWLLSP
```

```
HCV_K1_R3   KTSERSQPRG RRQPIPKVRR SEGRTWAQPG YPWPLYGNEG LGWAGWLLSP
HCV_K1_S1   KTSERSQPRG RRQPIPKARR PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
HCV_K1_S2   KTSERSQPRG RRQPIPKARQ PEGRAWAQPG YPWPLYGNEG MGWAGWLLSP
HCV_K1_S3   KTSERSQPRG RRQPIPKVRR SEGRTWAQPG YPWPLYGNEG LGWAGWLLSP
   HCV_L2   KTSERSQPRG RRQPIPKARQ PEGRAWAQPG YPWPLYANEG LGWAGWLLSP
    HCV_N   KTSERSQPRG RRQPIPKARR PEGRAWAQPG YPWPLYGNEG MGWAGWLLSP
 HCV12083   KTSERSQPRG RRQPIPKARQ PQGRHWAQPG YPWPLYGSEG CGWAGWLLSP
  HCV1480   KNSERSQPRG RRQPIPKARR PTGRSWGQPG YPWPLYANEG LGWAGWLLSP
 HCVPOLYP   KTSERSQPRG RRQPIPKARR PEGRAWAQPG YPWPLYGNEG MGWAGWLLSP
     HD_1   KTSERSQPRG RRQLIPKARQ PEGRSWAQPG YPWPLYGNEG MGWAGWLLSP
  HPCCGAA   KTSERSQPRG RRQPIPKARR PEGRTWAQPG YPWPLYGNEG CGWAGWLLSP
    HPCFG   KTSERSQPRG RRQPTPKARP REGRSWAQPG YPWPLYGNEG CGWAGWLLPP
HPCGENANTI  KTWERSQPRG RRQPIPKARQ PEGRAWAQPG YPWPLYGNEG LGWAGWLVSP
 HPCGENOM   KTSERSQPRG RRQPIPKARR PEGRTWAQPG YPWPLYGNEG FGWAGWLLSP
  HPCHUMR   KTSERSQPRG RRQPIPKARR PEGRTWAQPG YPWPLYGNEG LGWAGWLLSP
     HPCJ   KTSERSQPRG RRQPIPKARR PEGRTWAQPG YPWPLYGNEG MGWAGWLLSP
   HPCJCG   KTSERSQPRG RRQPIPKARR PEGRTWAQPG YPWPLYGNEG MGWAGWLLSP
 HPCJK046   KTSERSQPRG RRQPIPKARR QTGRAWGQPG YAWPLYGNEG CGWAGWLLSP
 HPCJK049   KTSERSQPRS RRQPIPRARR TEGRSWAQPG YPWPLYGNEG CGWAGWLLSP
   HPCJTA   KTSERSQPRG RRQPIPKARR PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
   HPCJTB   KTSERSQPRG RRQPIPKARR PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
   HPCK3A   KTSERSQPRG RRKPIPKARR SEGRSWAQPG YPWPLYGNEG CGWAGWLLSP
HPCPLYPRE   KTSERSQPRG RRQPIPKARR PEGRTWAQPG YPWPLYGNEG CGWAGWLLSP
  HPCPOLP   KTSERSQPRG RRQPIPKDRR STGKSWGKPG YPWPLYGNEG LGWAGWLLSP
    HPCPP   KTSERPQPRG RRQPIPKARQ PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
 HPCUNKCD   KTSERSQPRG RRQPIPKARR PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
    MKC1A   KTSERPQPRG RRQPIPKARQ PEGRAWAQPG YPWPLYGNEG LGWAGWLLSP
    NDM59   KTSERSQPRG RRQPIPKDRR STGKSWGKPG YPWPLYGNEG LGWAGWLLSP
     NZLI   KTSERSQPRG RRQPIPKARR SEGRSWAQPG YPWPLYGNEG CGWAGWLLSP
     SA13   KTSERSQPRG RRQPIPKARQ PTGRSWGQPG YPWPLYANEG LGWAGWLLSP
    Th580   KTSERSQPRG RRQPIPKARP SQGRTWGQPG YPWPLYGNEG CGWAGWLMSP
Type_3a_CB  KTSERSQPRR RRQPIPKARQ SGGRSWAQPG YPWPLYGNEG CGWAGWLLSP
   TypeV_D  KTSERSQPRG RRQPIPKARR SEGRSWAQPG YPWPLYGNEG CGWAGWLLSP
    VN004   KTSERSQPRG RRQPIPKARQ PIGRSWGQPG YPWPLYGNEG CGWAGWLLSP
    VN235   KTSERSQPRG RRQPIPKARR QTGRTWAQPG YPWPLYGNEG CGWMGWLLSP
    VN405   KTSERSQPRG RRQPIPKARQ SQGRHWAQPG YPWPLYGNEG CGWAGWLLSP

            101                                                 150
    BEBE1   RGSRPSWGPS DPRHKSRNLG KVIDTLTCGF ADLMGYIPVV GAPVGGVARA
    D89815  RGSRPSWGPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
ED43type_4  RGSRPSWGPN DPRGRSRNLG KVIDTLTCGF ADLMGYIPLV GAPVGSVARA
    HC_C2   RGSRPSWGPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
    HC_G9   RGSRPSWGPS DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
  HCU16326  RGSRPSWGPT DPRRKSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGVARA
 HCV_H_CMR  RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
    HCV_J1  RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
  HCV_J483  RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
    HCV_J8  RGSRPTWGPT DPRHRSRNLG RVIDTITCGF ADLMGYIPVV GAPVGGVARA
   HCV_JK1  YGSRPRWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
    HCV_JS  RGSRPNWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGVARA
 HCV_K1_R1  RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
 HCV_K1_R2  RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGASRA
 HCV_K1_R3  RGSRSSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPVV GPPLGGVARA
 HCV_K1_S1  RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGTARA
 HCV_K1_S2  RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGASRA
 HCV_K1_S3  RGSRSNWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPVV GPPLGGVARA
   HCV_L2   RGSRPSWGPT DPRRRSRNLG KVIDTPTCGF ADLMGYIPLV GAPLGGVARA
    HCV_N   RGSRPSWGPT DPRRRSRNLG KVIDTLTCGL ADLMGYVPLV GGPLGGAARA
 HCV12083   RGSRPHWGPN DPRRRSRNLG KVIDTLTCGF ADLMWYIPVV GAPLGGVAAA
  HCV1480   RSSRPNWGPN DPRRKSPNLG RVIHTLTCGF PHLMGYIPLV GGPVGGVSRA
 HCVPOLYP   RGSRPSWGPS DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
```

```
          HD_1  RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
        HPCCGAA RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
          HPCFG RGSRPSWGQN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLI GAPVGGVARA
     HPCGENANTI RGSRPNWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGVARA
       HPCGENOM RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
       HPCHUMR RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
          HPCJ RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
        HPCJCG RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
      HPCJK046 RGSRPTWGPN DPRRRSRNLG KVIDTLTCGL ADLMGYIPVI GGPLGGVAAA
      HPCJK049 RGSRPSWGPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPVGGVARA
        HPCJTA RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
        HPCJTB RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
        HPCK3A RGSRPNWAPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
     HPCPLYPRE RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
       HPCPOLP RGSRPSWGPN DPRHRSRNVG KVIDTLTCGF ADLMGYIPVV GAPLGGVARA
        HPCPP RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
      HPCUNKCD RGSRPSWGPT DPRRKSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGVARA
         MKC1A RGSRPSWGPT DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPLGGAARA
         NDM59 RGSRPSWGPT DPRHRSRNVG KVIDTLTCGF ADLMGYIPVV GAPLGGVARA
          NZLI RGSRPSWGPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPVGGVARA
          SA13 RGSRPNWGPN DPRRKSRNLG KVIDTLTCGF ADLMGYIPLV GGPVGGVARA
         Th580 RGSRPSWGPN DPRRRSRNLG KVIDTLTCGL ADLMGYIPVV GGPLGGVAAA
    Type_3a_CB RGSRPSWGPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPVGGVARA
       TypeV_D RGSRPSWGQN DPRRRSRNLG KVIDTLTCGF ADLMGYIPLV GAPVGGVARA
         VN004 RGSRPNWGPN DPRRRSRNLG KVIDTLTCGL ADLMGYIPVL GGPLGGVAAA
         VN235 RGSRPHWGPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPVV GAPLGGVAAA
         VN405 RGSRPNWGPN DPRRRSRNLG KVIDTLTCGF ADLMGYIPVV GAPLGGVAAA

               151                                                 200
          BEBE1 LAHGVRVLED GINYATGNLP GCSFSIFLLA LLSCISVPVS AVEVRNTSSS
          D89815 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVRNVSGI
    ED43type_4 LAHGVRALED GINYATGNLP GCSFSIFLLA LLSCLTVPAS AVNYRNVSGI
          HC_C2 LAHGVRALED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVRNASGV
          HC_G9 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTVPAS AVGVRNSSGV
        HCU16326 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTTPVS AYEVRNASGM
       HCV_H_CMR LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTVPAS AYQVRNSSGL
          HCV_J1 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTVPAS AYQVRNSTGL
        HCV_J483 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVRNVSGI
          HCV_J8 LAHGVRVLED GINYATGNLP GCSFSIFLLA LLSCVTVPVS AVEVRNISSS
         HCV_JK1 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTVPVS TYEVRNVSGV
          HCV_JS LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTTPAS AYEVRNVSGI
       HCV_K1_R1 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVCNASGL
       HCV_K1_R2 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVRNASGV
       HCV_K1_R3 LAHGVRVLEV GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVRNVSGV
       HCV_K1_S1 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLPCLTIPAS AYEVCNASGL
       HCV_K1_S2 LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVRNASGV
       HCV_K1_S3 LAHGVRVLEV GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVRNVSGV
          HCV_L2 LAHGVRVLED SVNYATGNLP GCSFSIFLLA LLSCLTVPAS AYEVRNVSGI
           HCV_N LAHGVRVLED GVNYATGNMP GCSFSIFLLA LLSCLTVPAS AHEVRNASGV
        HCV12083 LAHGVRAIED GINYATGNLP GCSFSIFLLA LLSCLTTPAS ALTYGNSSGL
         HCV1480 LAHGVKVLED GINYATGNLP GCPFSIFVLA LLWCLTVPAS AVPYRNASGV
        HCVPOLYP LAHGVRVLED GVNYATGNLP GCSFSIFLLG LLSCLTIPAS AYEVRNVSGV
          HD_1 LAHGVRVLEG GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVRNVSGV
        HPCCGAA LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTVPAS AYQVRNSSGL
          HPCFG LAHGVRALED GVNYATGNLP GCSFSIFLLA LFSCLTCPAS SLEYRNASGL
     HPCGENANTI LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVHNVSGI
       HPCGENOM LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTTPAS AYEVRNVSGI
       HPCHUMR LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTTPAS AYEVHNVSGI
          HPCJ LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTVPVS AYEVRNVSGG
        HPCJCG LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYEVRNVSGI
      HPCJK046 LAHGVRAVED GVNYATGNLP GCSFSIFLLA LLSCLTVPAS AVNYANKSGI
```

```
HPCJK049  LAHGVRALED GINFATGNLP GCSFSIFLLA LLSCLLTPTA GLEYRNASGL
  HPCJTA  LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYQVRNASGL
  HPCJTB  LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYQVRNRSGL
  HPCK3A  LAHGVRALED GINFATGNLP GCSFSIFLLA LFSCLIHPAA SLEWRNTSGL
HPCPLYPRE LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTVPAS AYQVRNSTGL
 HPCPOLP  LAHGVRVLED GVNFATGNLP GCSFSIFLLA LLSCITTPVS AAEVKNISTG
   HPCPP  LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYQVRNASGV
HPCUNKCD  LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTTPVS AYEVRNASGM
   MKC1A  LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCLTIPAS AYQVRNASGV
   NDM59  LAHGVRVLED GVNYATGNLP GCSFSIFLLA LLSCITVPVS AVQVKNISDS
    NZLI  LAHGVRALED GINFATGNLP GCSFSIFLLA LFSCLIHPAA SLEWRNTSGL
    SA13  LAHGVRVLED GVNYATGNLP GCSFSIFILA LLSCLTVPTS AVPYRNASGV
    Th580 LAHGVRAIED GINYATGNLP GCSFSIFILA LLSCLTTPAS ALTYGNSSGL
Type_3a_CB LAHGVRALED GINFATGNLP GCSFSIFLLA LFSCLIHPAA SLEWRNTSGL
  TypeV_D LAHGVRALED GINFATGNLP GCSFSIFLLA LFSCLIHPAA SLEWRNTSGL
    VN004 LAHGVRAIED GVNYATGNLP GCSFSIFLLA LLSCLTTPAS AIQVRNASGI
    VN235 LAHGVRAVED GINYATGNLP GCSFSIFLLA LLSCLTTPAS AVHYANKSGI
    VN405 LAHGVRAIED GINYATGNLP GCSFSIFLLA LLSCLTTPAS AVHYRNISGI

          201                                                 250
    BEBE1 YMATNDCSNS SIVWQLEGAV LHTPGCVPCE KTGN.KSRCW VPVTPNIAIN
    D89815 YHVTNDCSNS SIVYEAADVI MHAPGCVPCV RENN.SSRCW VALTPTLAAR
ED43type_4 YHVTNDCPNS SIVYEADHHI MHLPGCVPCV REGN.QSRCW VALTPTVAAP
    HC_C2 YHVTNDCSNS SIVYEAADMI MHNPGCVPCV RENN.SSRCW VALTPTLAAR
    HC_G9 YHVTNDCPNA SVVYETENLI MHLPGCVPYV REGN.ASRCW VSLSPTVAAR
  HCU16326 YHVTNDCSNS SIVYEAADMI MHTPGCVPCV REDN.SSRCW VALTPTLAAR
 HCV_H_CMR YHVTNDCPNS SIVYEAADAI LHTPGCVPCV REGN.ASRCW VAVTPTVATR
   HCV_J1 YHVTNDCPNS SIVYEAHDAI LHTPGCVPCV REGN.VSRCW VAMTPTVATR
 HCV_J483 YHVTNDCSNS SIVYEAADMI MHTPGCVPCV REDN.SSRCW VALTPTLAAR
   HCV_J8 YYATNDCSNN SITWQLTDAV LHLPGCVPCE NDNG.TLHCW IQVTPNVAVK
  HCV_JK1 YHVTNDCSNS SIVYEAADMI MHTPGCVPCV REGN.SSRCW VALTPTLAAR
   HCV_JS YHVTNDCSNS SIVYEAADII MHTPGCVPCV REKN.ISRCW VALTPTLAAR
 HCV_K1_R1 YHVTNDCSNS SIVYEAQDMI MHTPGCVPCV RENN.SSRCW VALTPTLAAR
 HCV_K1_R2 YHVTNDCSNA SIVYEAADMI MHTPGCVPCV REAN.SSRCW VALTPTLAAR
 HCV_K1_R3 YHVTNDCSNS SIVYEAEDVI MHTPGCLPCV RENN.SSRCW VALTPTLAGR
 HCV_K1_S1 YHVTNDCSNS SIVYEAEDMI MHTPGCVPCV RENN.SSRCW VALTPTLAAR
 HCV_K1_S2 YHVTNDCSNA SIVYEAADMI MHTPGCVPCV REAN.SSRCW VALTPTLAAR
 HCV_K1_S3 YHVTNDCSNS SIVYEAEDVI MHTPGCLPCV RENN.SSRCW VALTPTLAGR
   HCV_L2 YHVTNDCSNS SIVYEAADLI MHTPGCVPCV REAN.SSRCW VALTPTLAAR
    HCV_N YHVTNDCSNS SIVFEAADLI MRTPGCVPCV REGN.SSRCW VALTPTLAAR
  HCV12083 YHLTNDCSNS SIVLEADAMI LHLPGCLPCV RVGN.QSTCW HAVSPTLATP
  HCV1480 YHVTNDCPNS SIVYEADNLI LHAPGCVPCV LEDN.VSRCW VQITPTLSAP
 HCVPOLYP YHVTNDCSNA SIVYEAADMI MHVPGCVPCV RVDN.SSRCW VALTPTLAAR
     HD_1 YHVTNDCSNS SIVYETADMI MHTPGCVPCV REDN.SSRCW VALTPTLAAR
  HPCCGAA YHVTNDCPNS SVVYEAADAI LHTPGCVPCV REGN.ASRCW VAVTPTVATR
    HPCFG YLLTNDCSNR SIVYEADDVI LHLPGCVPCV ETDNNNTSCW TPISPTVAVK
HPCGENANTI YHVTNDCSNS SIVYEAADMI MHTPGCVPCV RENN.SSRCW VALTPTLAAR
 HPCGENOM YHVTNDCSNS SIVYEAADLI MHTPGCVPCV REGN.SSRCW VALTPTLAAR
  HPCHUMR YHVTNDCSNA SIVYEAADLI MHTPGCVPCV REGN.SSRCW VALTPTLAAR
    HPCJ  YHVTNDCSNS SIVYEAADVI MHTPGCVPCV RENN.SSRCW VALTPTLAAR
   HPCJCG YHVTNDCSNS SIVYEAADMI MHTPGCVPCV RESN.FSRCW VALTPTLAAR
  HPCJK046 YHLTNDCPNS SMVYEAEAII LHLPGCVPCI RTGN.QSRCW TPATPTLAIP
  HPCJK049 YTVTNDCSNG SIVYEAGDVI LHLPGCIPCV RLNN.ASKCW TPVSPTVAVS
   HPCJTA YHVTNDCSNS SIVYEAAGMI MHTPGCVPCV RENN.ASRCW VALTPTLAAR
   HPCJTB YHVTNDCSNS SIVYEAAGMI MHTPGCVPCV RENN.VSRCW VALTPTLAAR
   HPCK3A YVLTNDCSNS SIVYEADDVI LHTPGCIPCV QDGN.TSTCW TPVTPTVAVR
HPCPLYPRE YHVTNDCPNS SIVYEAADAI LHTPGCVPCV REGN.ASRCW VAMTPTVATR
  HPCPOLP YMVTNDCTND SITWQLQAAV LHVPGCVPCE KVGN.TSRCW IPVSPNVAVQ
   HPCPP YHVTNDCSNS SIVYEAADVI MHTPGCVPCV RENN.SSRCW VALTPTLAAR
HPCUNKCD YHVTNDCSNS SIVYEAADMI MHTPGCVPCV REDN.SSRCW VALTPTLAAR
   MKC1A YHVTNDCSNS SIVYEAADVI MHTPGCVPCV RENN.SSRCW VALTPTLAAR
```

```
      NDM59  YMVTNDCSND SITWQLQAAV LHVPGCVPCE KMGN.ISRCW IPVSPNVAVQ
       NZLI  YVLTNDCSNS SIVYEADDVI LHTPGCVPCV QDGN.TSTCW TPVTPTVAVR
       SA13  YHVTNDCPNS SIVYEAEDLI LHAPGCVPCV RQGN.VSRCW VQITPTLSAP
      Th580  YHLTNDCPRS SIVLEAEAMI LHLAGCVPCV RAGN.ISRCW HPVSPTLAVP
  Type_3a_CB  YVLTNDCSNS SIVYEADDVI LHTPGCVPCV QNDN.ISTCW TPVTPTVAVR
   TypeV_D  YVLTNDCSNS SIVYEADDVI LHTPGCVPCV QDGN.TSTCW TPVTPTVAVR
      VN004  YHLTNDCSNN SIVFEAETII LHLPGCVPCI KVGN.GSRCW LSVSPTLAVP
      VN235  YHLTNDCPNS SIVYEAEDFI MHLPGCVPCI KSGN.GSSCW LPATLTIAVP
      VN405  YHLTNDCPNS SIIYEADNII MHTPGCVPCV KTGN.KSQCW VPVAPTLAVA

                251                                              300
      BEBE1  QPGALTKGLR AHIDVIVMSA TLCSALYVGD VCGALMIAAQ VVVVSPQHHH
      D89815  NASVPTTTLR RHVDLLVGTA AFCSAMYVGD LCGSVFLISQ LFTFSPRRHE
  ED43type_4  YIGAPLESLR SHVDLMVGAA TVCSGLYIGD LCGGLFLVGQ MFSFRPRRHW
      HC_C2  NASVPTTTIR RHVDLLVGAA ALCSAMYVGD LCGSVFLVSQ LFTFSPRRHE
      HC_G9  DSRVPVSEVR RRVDSIVGAA AFCSAMYVGD LCGSIFLVGQ IFTFSPRHHW
    HCU16326  NASVPTTTLR RHVDLLVGVA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRHE
   HCV_H_CMR  DGKLPTTQLR RHIDLLVGSA TLCSALYVGD LCGSVFLVGQ LFTFSPRRHW
      HCV_J1  DGKLPATQLR RHIDLLVGSA TLCSALYVGD LCGSVFLIGQ LFTFSPRRHW
    HCV_J483  NASVPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRHE
      HCV_J8  HRGALTRSLR THVDMIVMAA TACSALYVGD VCGAVMILSQ AFMVSPQRHN
     HCV_JK1  NSSIPTTTIR RHVDLLVGAA ALCSAMYVGD LCGSVFLVSQ LFTFSPRRYE
      HCV_JS  NISVPTATIR RHVDLLVGTA AFCSAMYVGD LCGSVFLVSQ LFTFSPRWHE
   HCV_K1_R1  NASIPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRYE
   HCV_K1_R2  NSSVPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRHE
   HCV_K1_R3  NSSIPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRYE
   HCV_K1_S1  NASIPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRYE
   HCV_K1_S2  NSSVPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRHE
   HCV_K1_S3  NSSIPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFAFSPRRYE
      HCV_L2  DSSIPTATIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRLHQ
       HCV_N  NATIHTTTIR HHVDLLVGAA ALCSAMYVGD LCGSVFLVSQ LFTFSPRRHA
    HCV12083  NASTPATGFR RHVDLLAGAA VVCSSLYIGD LCGSLFLAGQ LFAFQPRRHW
     HCV1480  SFGAVTALLR RAVDYLAGGA AFCSALYVGD ACGALSLVGQ MFTYKPRQHT
    HCVPOLYP  NASVPTTAIR RHVDLLVGAA TFCSAMYVGD LCGSVFLVAQ LFTFSPRRHE
        HD_1  NASIPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRHA
     HPCCGAA  DGKLPTTQLR RHIDLLVGSA TLCSALYVGD LCGSVFLVGQ LFTFSPRHHW
       HPCFG  HPGVTTASIR NHVNMLVAPP TLCSALYVED AFGAVSLVGQ AFTFRPRQHK
  HPCGENANTI  NNSVPTATIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRYE
    HPCGENOM  NATIPTATVR RHVDLLVGAA AFSSAMYVGD LCGSVFLVSQ LFTFSPRRYE
     HPCHUMR  NVTIPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRHV
        HPCJ  NTTVPTTTIR RHVDLLVGTA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRHE
     HPCJCG  NSSIPTTTIR RHVDLLVGAA ALCSAMYVGD LCGSVFLVSQ LFTFSPRRYE
    HPCJK046  NSTVPASGFR QHIDLMVGAA ALCSAMYLGD LCGGVFLVGQ LFTFRPRIHQ
    HPCJK049  RPGAATASLR THVDMMVGAA TLCSALYVGD LCGALFLVGQ GFSWRHRQHW
     HPCJTA  NTSIPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRYE
     HPCJTB  NTSIPTTTIR RHVDLLVGAA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRYE
     HPCK3A  YVGATTASIR SHVDLLVGAG TMCSALYVGD MCGPVFLVGQ AFTFRPRRHR
   HPCPLYPRE  DGKLPATQLR RHIDLLVGSA TLCSALYVGD LCGSVFLVGQ LFTFSPRRHW
     HPCPOLP  QPGALTQGLR THIDMVVMSA TLCSALYVGD LCGGVMLAAQ MFIVSPQHHW
       HPCPP  NSSIPTTTIR RHVDLLVGAA ALCSAMYVGD FCGSVFLVSQ LFTFSPRRYE
    HPCUNKCD  NASVPTTTLR RHVDLLVGVA AFCSAMYVGD LCGSVFLVSQ LFTFSPRRHE
      MKC1A  NSSIPTTTIR RHVDLLVGAA ALCSAMYVGD FCGSVFLVSQ LFTFSPRRYE
      NDM59  RPGALTQGLR AHIDMVVMSA TLCSALYVGD LCGGVMLAAQ MFIVSPQHHH
       NZLI  YVGATTASIR SHVDLLVGAA TMCSALYVGD MCGAVFLVGQ AFTFRPRRHQ
       SA13  SLGAVTAPLR RAVDYLAGGA ALCSALYVGD ACGAVFLVGQ MFTYSPRRHN
      Th580  NASVPASGFR KHVDLLAGAA VVCSSMYIGD LCGAVFLAGQ LATFSPRIHD
  Type_3a_CB  YVGATTASIR SHVDLLVGAA TMCSALYVGD MCGAVFLVGQ AFTFRPRRHQ
   TypeV_D  YVGATTASIR SHVDLLVGAA TMCSALYVGD MCGAVFLVGQ AFTFRPRRHQ
      VN004  NSSVPIHGFR RHVDLLVGAA AFCSAMYIGD LCGSVFLVGQ LFTFRPKHHQ
      VN235  NASIPVRGFR RHVDLMVGAA AFCSAMYVGD LCGGIFLVGQ LFSFNPRRHW
      VN405  NASVPIRGFR SHVDLLVGSA AACSALYIGD LCGGVFLVGQ LFTFRPRQHT
```

	301				350
BEBE1	FVQECNCSIY	PGKITGHRMA	WDMMMNWSPT	TTMLLAYLVR	IPEVVLDIIT
D89815	TVQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	AALVVSQLLR	IPQAVMDMVA
ED43type_4	TTQDCNCSIY	TGHITGHRMA	WDMMMNWSPT	TTLVLAQVMR	IPTTLVDLLS
HC_C2	TVQDCNCSIY	PGHITGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVMDMVA
HC_G9	TTQDCNCSIY	PGHVTGHRMA	WDMMMNWSPT	GALVVAQLLR	IPQAIVDMIA
HCU16326	TVQDCNCSIY	PGRVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVT
HCV_H_CMR	TTQSCNCSIY	PGHITGHRMA	WDMMMNWSPT	AALVVAQLLR	IPQAIMDMIA
HCV_J1	TTQGCNCSIY	PGHITGHRMA	WDMMMNWSPT	AALVMAQLLR	IPQAILDMIA
HCV_J483	TVQDCNCSIY	PGHLSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVA
HCV_J8	FTQECNCSIY	QGHITGHRMA	WDMMLSWSPT	LTMILAYAAR	VPELVLEIIF
HCV_JK1	TVQDCNCSLY	PGHVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVV
HCV_JS	TVQDCNCSLY	PGHVSGHRMA	WDMMMNWSPT	AALVVSQLLR	IPQAVVDMVA
HCV_K1_R1	TVQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVA
HCV_K1_R2	TVQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVMDMVA
HCV_K1_R3	TVQDCNCSLY	PGHITGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVA
HCV_K1_S1	TVQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVA
HCV_K1_S2	TVQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVMDMVA
HCV_K1_S3	TVQDCNCSLY	PGHITGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVMDMVA
HCV_L2	TVQDCNCSIY	PGHLTGHRMA	WDMMMNWSPT	AALVVSQLLR	IPQAIVDMVA
HCV_N	TLQDCNCSIY	PGHASGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVIDMVA
HCV12083	TVQDCNCSIY	TGHVTGHKMA	WDMMMNWSPT	TTLVLSSILR	VPEICASVIF
HCV1480	TVQDCNCSIY	SGHITGHRMA	WDMMMKWSPT	TALLMAQLLR	IPQVVIDIIA
HCVPOLYP	TVQDCNCSIY	PGHITGHRMA	WDMMMNWSPT	AALVVSQLLR	IPQAIVDMVA
HD_1	TVQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	AALVVSQLLR	IPQAVVDMVA
HPCCGAA	TTQDCNCSIY	PGHITGHRMA	WNMMMNWSPT	AALVVAQLLR	IPQAIMDMIA
HPCFG	TVQTCNCSIY	PGHVSGHRMA	WDMMMNWSPA	IGLVISHLMR	LPQTFFDLVV
HPCGENANTI	TVQDCNCSIY	PGHVTGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVG
HPCGENOM	TIQDCNCSIY	PGHVTGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVMDMVV
HPCHUMR	TLQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVA
HPCJ	TVQDCNCSIY	PGHLSGHRMA	WDMMMNWSPT	AALVVSQLLR	IPQAVVDMVA
HPCJCG	TVQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVA
HPCJK046	TVQDCNCSIY	TGHVTGHRMA	WDMMMNWSPT	ATFVVSSALR	APQVLFDIFA
HPCJK049	TVQDCNCSIY	PGHLTGHRMA	WDMMMNWSPA	MTLIVSQVLR	LPQTMFDLVI
HPCJTA	TVQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVA
HPCJTB	TVQDCNCSIY	PGHVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVA
HPCK3A	TVQTCNCSLY	PGHLSGQRMA	WDMMMNWSPA	VGMVVAHILR	LPQTLFDVVA
HPCPLYPRE	TTQGCNCSIY	PGHITGHRMA	WDMMMNWSPT	TALVMAQLLR	IPQAILDMIA
HPCPOLP	FVQDCNCSIY	PGTITGHRMA	WDMMMNWSPT	ATMILAYAMR	VPEVIIDIIG
HPCPP	TVQDCNCSIY	PGHVSGHRMA	WDMIMNWSPT	TALVVSQLLR	IPQAVVDMVA
HPCUNKCD	TVQDCNCSIY	PGRVSGHRMA	WDMMMNWSPT	TALVVSQLLR	IPQAVVDMVT
MKC1A	TVQDCNCSIY	PGHVSGHRMA	WDMIMNWSPT	TALVVSQLLR	IPQAVVDMVA
NDM59	FVQECNCSIY	PGAITGHRMA	WDMMMNWSPT	ATMILAYAMR	VPEVIIDIIS
NZLI	TVQTCNCSLY	PGHLSGHRMA	WDMMMNWSPA	VGMVVAHVLR	LPQTLFDIMA
SA13	VVQDCNCSIY	SGHITGHRMA	WDMMMNWSPT	TALVMAQLLR	IPQVVIDIIA
Th580	ITQDCNCSVY	TGHVTGHRMA	WDMMMNWSPT	TTLVLSSILR	VPEIVLEVFA
Type_3a_CB	TVQTCNCSLY	PGHLSGHRMA	WDMMMNWFPA	LGMAVAHVLR	VPQTLFDIIA
TypeV_D	TVQTCNCSLY	PGHLSGHRMA	WDMMMNWSPA	VGMVVSHVLR	LPQTLFDIIA
VN004	VTQDCNCSIY	AGHITGHRMA	WDMMLNWSPT	VSYVVSSALR	VPQLLLEVIT
VN235	VVQDCNCSIY	VGHITGHRMA	WDMMMNWSPT	ATLVLSYVMR	IPQVIMDIFT
VN405	TVQECNCSIY	TGHITGHRMA	WDMMMNWSPT	VTFITSSLLR	VPQLLLEIAL

```
            351                                                        400
     BEBE1  GGHWGVMFGL AYFSMQGAWA KVVVILLLTA GVEASTYTTG AVVGRSTHLF
     D89815 GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGHTRVTG GVQGHVTSTL
ED43type_4 GGHWGVLVGV AYFSMQANWA KVILVLFLFA GVDAETHVSG AAVGRSTAGL
      HC_C2 GAHWGVLAGL AYYSMVGNWA KVLIVLLLFA GVDGNTYVTG GAAARGASGI
      HC_G9 GAHWGVLAGL AYYSMVGNWA KVVVVLLLFA GVDAETRVTG GAAGHTAFGF
   HCU16326 GSHWGILAGL AYYSMVGNWA KVLIAMLLFA GVDGTTHVTG GAQGRAASSL
  HCV_H_CMR GAHWGVLAGI AYFSMVGNWA KVLVVLLLFA GVDAETHVTG GSAGHTTAGL
      HCV_J1 GAHWGVLAGI AYFSMVGNWA KVLVVLLLFA GVDAETIVSG GQAARAMSGL
   HCV_J483 GAHWGVLAGL AYYSMVGNWA KVLIVALLFA GVDGETYTSG GAASHTTSTL
      HCV_J8 GGHWGVVFGL AYFSMQGAWA KVIAILLLVA GVDATTYSSG QEAGRTVAGF
     HCV_JK1 GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGTTYVSV GHASQTTRRV
      HCV_JS GAHWGVLAGL AYYSMIGNWA KVLIVMLLFA GADGTTHVTG GVQAHGAYGL
   HCV_K1_R1 GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGNTYTTG AAQGRTVSRL
   HCV_K1_R2 GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGNTYVSG GAKSHTTQGL
   HCV_K1_R3 GAHWGVLAGL VYYSMVGNWA KVLIVMLLFV GVDGSTHVSG GTTAYNTRSF
   HCV_K1_S1 GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGDTYTTG GAQGHTTSRV
   HCV_K1_S2 GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGRTTVTG GAQGHTTQRL
   HCV_K1_S3 GAHWGVLAGL VYYSMVGNWA KVLIVMLLFV GVDGSTRVSG GTTAYNTRGL
      HCV_L2 GAHWGVLAGL AYYPMVGNWA KVLIVMLLFA GVDGTTVTMG GTVARTTYGF
       HCV_N GAHWGVLAGL AYYSMAGNWA KVLIVMLLFA GVDGHTLTTG GHAAHLTSGF
   HCV12083 GGHWGILLAV AYFGMAGNWL KVLAVLFLFA GVEAQT.MIA HGVSQTTSGF
    HCV1480 GGHWGVLLAA AYFASTANWA KVILVLFLFA GVDGRTHTVG GTVGQGLKSL
  HCVPOLYP  GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGDTHTTG GVAGRDTLRF
       HD_1 GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGTTTVTG GSQARTVYEL
    HPCCGAA GAHWGVLAGI KYFSMVGNWA KVLVVLLLFA GVDAETHVTG GNAGRTTAGL
      HPCFG GAHWGVMAGL AYFSMQGNWA KVVIVLIMFS GVDATTHTTG GSAAQATAGF
 HPCGENANTI GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGSTIVSG GTVARTTHSL
   HPCGENOM GAHWGVLAGL AYYAMVGNWA KVLIVMLLFA GVDGDTYASG GAQGRSTLGF
    HPCHUMR GAHWGVLAGL AYYSMAGNWA KVLIVMLLFA GVDGDTHVTG GAQAKTTNRL
       HPCJ GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGQTYTTG GAVARTTTGF
     HPCJCG GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGHTHVTG GRVASSTQSL
   HPCJK046 GGHWGIIGAL LYYSTAANWA KVIIVLLLFA GVDAST.YVA SSVSQATSGL
   HPCJK049 GAHWGVMAGV AYYSMQGNWA KVFLVLCLFS GVDASTTITG GVAASGAFTI
     HPCJTA GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGVTYTTG GSQARHTQSV
     HPCJTB GAHWGVLAGL AYYSMVGNWA KVLIVMLLFA GVDGVTYTTG GSQARHTQGV
     HPCK3A GAHWGIIAGL AYYSMQGNWA KVAIIMVMFS GVDASTHVTA GQAARNAYGI
  HPCPLYPRE GAHWGVLAGI AYFSMVGNWA KVLVVLLLFA GVDAETHVTG GSAGHTVSGF
     HPCPOLP GAHWGVMFGL AYFSMQGAWA KVVVILLLAA GVDAQTHTVG GSTAHNARTL
      HPCPP GAHWGVLAGL AYYSMVGNWA KVLVVMLLFA GVDGGTHVTG GKVAYTTQGF
   HPCUNKCD GSHWGILAGL AYYSMVGNWA KVLIAMLLFA GVDGTTHVTG GAQGRAASSL
      MKC1A GAHWGVLAGL AYYSMVGNWA KVLVVMLLFA GVDGGTHVTG GKVAYTTQGF
      NDM59 GAHWGVMFGL AYFSMQGAWA KVVVILLLTA GVDAHTRSIA GSVAHATSGL
       NZLI GAHWGILAGL AYYSMQGNWA KVAIIMVMFS GVDAHTYTTG GTASRHTQAF
       SA13 GAHWGVLFAA AYYASAANWA KVVLVLFLFA GVDANTRTVG GSAAQGARGL
       Th580 GGHWGVLIAI AYFGMSGNWL KVIAVLFLFA GVEATT.TVG RAAGRSAYLF
  Type_3a_CB GAHWGILAGL AYYSMQGNWA KVAIIMVMFS GVDAVTYTTG GSAAHATRGL
    TypeV_D GAHWGILAGL AYYSMQGNWA KVAVIMVMFS GVDAETYITG GSAAHGVSTL
      VN004 GAHWGVLGAL LYFSMVANWA KVIAVLFLFA GADATT.YTG SAVSSTTGAF
      VN235 GGHWGILAGI LYYSMVANWA KVLCILFLFA GVDATTRTTG AQAARATLGF
      VN405 EGHWGVIGAL LYYSMVANWA KVFAVLLLFA GVDATT.HIG SSASATTNRL

            401                                                        450
     BEBE1 TSMFSLGSQQ RVQLIHTNGS WHINRTALNC NDSLETGFLA ALFYTSSFNS
     D89815 TSLFRPGASQ KIQLVNTNGS WHINRTALNC NDSLKTGFLA ALFYTHKFNA
ED43type_4 ANLFSSGSKQ NLQLINSNGS WHINRTALNC NDSLNTGFLA SLFYTHKFNS
      HC_C2 TSLFSRGPSQ KIQLVNTNGS WHINRTALNC NDSFNTGFLA ALFYAHRFNS
      HC_G9 ASFLAPGAKQ KIQLINTNGS WHINRTALNC NESLDTGWLA GLLYYHKFNS
   HCU16326 TSLFSPGPVQ HLQLINTNGS WHINRTALSC NDSLNTGFVA ALFYKYRFNA
  HCV_H_CMR VGLLTPGAKQ NIQLINTNGS WHINSTALNC NDSLTTGWLA GLFYRHKFNS
      HCV_J1 VSLFTPGAKQ NIQLINTNGS WHINSTALNC NESLNTGWLA GLIYQHKFNS
```

```
HCV_J483      ASLFSPGASQ RIQLVNTNGS WHINRTALNC NDSLHTGFLA ALFYTHRFNS
HCV_J8        AGLFTTGAKQ NLYLINTNGS WHINRTALNC NDSLQTGFLA SLFYTHKFNS
HCV_JK1       ASFFSPGSAQ KIQLVNTNGS WHINRTALNC NESINTGFFA ALFYVKKFNS
HCV_JS        ASLFNVGPHQ KIQLVNTNGS WHINRTALNC NDTLQTGFLA ALFYKHRFNA
HCV_K1_R1     TSIFTPGASQ NIQLINTNGS WHINRTALNC NDSLQTGFIA ALFYARRFNS
HCV_K1_R2     VSLFAPGAQQ KIQLVNTNGS WHINRTALNC NDSLNTGFLA ALFYTHRFNS
HCV_K1_R3     TSLFSSGAQQ KIQLVNTNGS WHINRTALNC NDTLQTGFLA ALFYTHRFND
HCV_K1_S1     ASLFTSGASQ NIQLINTNGS WHINRTALNC NDSLQTGFIA ALFYAHRFNS
HCV_K1_S2     ASLFTFGAQQ RIQLVNTNGS WHINRTALNC NDSLNTGFLA ALFYTHRFNS
HCV_K1_S3     TSLFSSGAQQ KIQLVNTNGS WHINRTALNC NDTLQTGFLA ALFYTHRFND
HCV_L2        TGLFRPGASQ KIQLINTNGS WHINRTALNC NDSLNTGFLA ALFYTHRFNA
HCV_N         AGLFTPGPSQ RIQLINTNGS WHINRTALNC NDSLQTGFLA ALSYTYRFNS
HCV12083      ASLLTPGAKQ NIQLINTNGS WHINRTALNC NDSLQTGFLA SLFYTHKFNS
HCV1480       TSFFNPGPQR QLQFVNTNGS WHINSTALNC NDSLQTGFIA GLMYAHKFNS
HCVPOLYP      TGFFSLGPKQ KIQLVNTNGS WHINRTALNC NDSLNTGWLA ALFYTHSFNA
HD_1          TSLYTRGPSQ RIQLVNTNGS WHINRTALNC NDSLQTGFLA ALFYTRSFNS
HPCCGAA       VGLLTPGAKQ NIQLINTNGS WHINSTALNC NESLNTGWLA GLFYQHKFNS
HPCFG         TSFFTRGPSQ NLQLVNSNGS WHINSTALNC NDSLNTGFIA GLFYYHKFNS
HPCGENANTI    ASLFTQGASQ KIQLINTNGS WHINRTALNC NDSLQTGFLA SLFYAHRFNA
HPCGENOM      TSLFTPGASQ KIQLINTNGS WHINRTALNC NDSLNTGFLA ALFYTHRFNA
HPCHUMR       VSMFASGPSQ KIQLINTNGS WHINRTALNC NDSLQTGFLA ALFYTHSFNS
HPCJ          ASLFSAGSQE NIQLINTNGS WHINRTALNC NDSLNTGFLA ALFYTHKFNS
HPCJCG        VSWLSQGPSQ KIQLVNTNGS WHINRTALNC NDSLQTGFIA ALFYAHRFNA
HPCJK046      VSLFSAGARQ NLQLINTNGS WHINRTALNC NDSLQTGFIA SLFYRNKFNA
HPCJK049      TSLFSTGAKQ PLHLVNTNGS WHINRTALNC NDSLNTGFIA GLLYYHKFNS
HPCJTA        TSFFTQGPAQ RIQLINTNGS WHINRTALNC NESLNTGFFA ALFYAHKFNS
HPCJTB        ASFFTPGPAQ KIQLINTNGS WHINRTALNC NESLNTGFFA ALFYAHKFNS
HPCK3A        TSLFSVGAKQ NLQLINTNGS WHINRTALNC NESINTGFIA GLFYYHKFNS
HPCPLYPRE     VSLLAPGAKQ NVQLINTNGS WHLNSTALNC NDSLNTGWLA GLFYHHKFNS
HPCPOLP       TGMFSLGARQ KIQLINTNGS WHINRTALNC NDSLHTGFLA SLFYTHSFNS
HPCPP         TSFFSRGPSQ KIQLVNTNGS WHINRTALNC NDSLNTGFLA ALFYTHSFNA
HPCUNKCD      TSLFSPGPVQ HLQLINTNGS WHINRTALSC NDSLNTGFVA ALFYKYRFNA
MKC1A         TSFFSRGPSQ KIQLVNTNGS WHINRTALNC NDSLNTGFLA ALFYTHSFNA
NDM59         AGLFTSGAKQ NIQLINTNGS WHINRTALNC NDSLNTGFIA SLFYTYRFNS
NZLI          AGLFDIGPQQ KLQLVNTNGS WHINSTALNC NESINTGFIA GLFYYHKFNS
SA13          ASLFTPGPQQ NLQLINTNGS WHINRTALNC NDSLQTGFVA GLLYYHKFNS
Th580         TSIFSSGPNQ KIQLINTNGS WHINRTALNC IDSLQTGFLS ALFYRSNFNS
Type_3a_CB    TSLFSVGAQQ KLQLVNTNGS WHINSTALNC NESINTGFIA GLFYYHRFNS
TypeV_D       TSLFSSGPQQ KLQLVKTNGS WHINSTALNC NESINTGFIA GLFYYHKFNS
VN004         VSLFSPGPTQ NLQLVNSNGS WHINRTALNC NDSLQTGFIA GLFARYKFNS
VN235         TGLFQTGAKQ NIHLINTNGS WHINRTALNC NDSLNTGFMA ALFYLHKFNS
VN405         TSFFSPGSKQ NVQLIKTNGS WHINRTALNC NDSLHTGFIA GLLYAHRFNS

              451                                                 500
BEBE1         SGCPERLAAC RSIESFRIGW GSLEYEESVT NDADMRPYCW HYPPRPCGIV
D89815        SGCPERMASC RSIDKFDQGW GPITYAQPD. .NSDQRPYCW HYAPRQCGIV
ED43type_4    SGCSERLACC KSLDSYGQGW GPLGVANISG .SSDDRPYCW HYAPRPCGIV
HC_C2         SGCPERMASC RSIDKFDQGW GPITYYQGD. .SPDQRPYCW HYPPRPCGIV
HC_G9         SGCPERMASC QPLTAFDQGW GPITHEGNA. .SDDQRPYCW HYALRPCGIV
HCU16326      SGCPERLATC RPIDTFAQGW GPITYTEPH. .DLDQRPYCW HYAPQPCGIV
HCV_H_CMR     SGCPERLASC RRLTDFAQGW GPISYANGS. .GLDERPYCW HYPPRPCGIV
HCV_J1        SGCPERLASC RRLTDFDQGW GPISHANGS. .GPDQRPYCW HYPPKPCGIV
HCV_J483      SGCPERMASC RPIDWFAQGW GPITYTEPD. .SPDQRPYCW HYAPRPCGIV
HCV_J8        SGCPERLSSC RGLDDFRIGW GTLEYETNVT NDGDMRPYCW HYPPRPCGIV
HCV_JK1       SGCSERMASC RPIDRFAQGW GPITHAESR. .SSDQRPYCW HYAPQPCGIV
HCV_JS        SGCPERMASC RPIDKFAQGW GPITYAEPD. .RLDQRPYCW HYPPRPCGIV
HCV_K1_R1     SGCPQRLASC RSIDKFAQGW GPITYAEGH. .DSDQRPYCW HYAPRPCGIV
HCV_K1_R2     SGCPERIASC RSIDAFAQGW GPITYAEPG. .SSDQRPYCW HYAPRPCGIV
HCV_K1_R3     SGCPERMASC RPIDKFAQGW GPITYVKPD. .ILDQRPYCW HYAPRPCGIV
HCV_K1_S1     SGCPQRLASC RSIDKFAQGW GPITYTEGH. .NSDQRPYCW HYAPRPCGIV
HCV_K1_S2     SGCPERIASC RSIDAFAQGW GPITYAEPG. .SSDQRPYCW HYAPRPCGIV
```

```
HCV_K1_S3   SGCPERMASC RPIDKFAQGW GPITYAKPD. .ILDQRPYCW HYAPRPCGIV
  HCV_L2    SGCPERMASC QSIDKFVQGW GPITYAENG. .SSDQRPYCW HYAPRRCGIV
   HCV_N    SGCPGRMASC RSIDKFDQGW GPITYADPK. .DPDQRPYCW HYAPQQCGII
 HCV12083   SGCPERMAAC KPLAEFRQGW GQITHKN.VS GPSDDRPYCW HYAPRPCEVV
 HCV1480    SGCPERMSSC RPLAAFDQGW GTISYATISG .PSDDKPYCW HYPPRPCGVV
 HCVPOLYP   SGCPERMASC HPIDEFAQGW GPITYAEHS. .SSDQRPYCW HYAPQPCGIV
   HD_1     SGCPERMASC RSIDQFDQGW GPITYAEPR. .DLDQRPYCW HYAPRPCGIV
 HPCCGAA    SGCPERLASC RRLTDFAQGW GPISYANGS. .GLDERPYCW HYPPRPCGIV
  HPCFG     SGCPERMSSC KPITYFNQGW GPLTDANING .PSEDRPYCW HYPPRPCNIT
HPCGENANTI  SGCPERMASC RSIDKFDQGW GPITYTEAD. .IQDQRPYCW HYAPRPCGIV
 HPCGENOM   SGCAERMASC RPIDTFDQGW GPITYTEPD. .SSDQRPYCW HYAPRKCGIV
 HPCHUMR    SGCPERMAQC RTIDKFDQGW GPITYAESS. .RSDQRPYCW HYPPPQCTIV
  HPCJ      SRAESVLASC RFIDEFDQGW GPITYTERN. .SSDQRPYCW HYPPRQCGII
 HPCJCG     SGCPERMASC RPIDEFAQGW GPITHDMPE. .SSDQRPYCW HYAPRPCGIV
 HPCJK046   TGCPERLSAC KTLDSFDQGW GPITYAN.IS GPAVEKPYCW HYPPRPCEVV
 HPCJK049   SGCVERMSAC SPLDRFAQGW GPLGPANISG .PSSEKPYCW HYAPRPCDTV
  HPCJTA    SGCPERMASC SSIDKFAQGW GPITYTEPR. .DLDQRPYCW HYAPRQCGIV
  HPCJTB    SGCPERMASC SSIDKFAQGW GPITYTEPG. .DLDQRPYCW HYAPRQCGIV
  HPCK3A    TGCPQRLSSC KPITFFKQGW GPLTDANITG .PSDDKPYCW HYAPRPCGIV
HPCPLYPRE   SGCPERLASC RPLTDFDQGW GPISYANGS. .GPDQRPYCW HYPPKPCGIV
 HPCPOLP    SGCPERMSAC RSIEAFRVGW GALQYEDNVT NPEDMRPYCW HYPPRQCGVV
  HPCPP     SGCPERMAGC RPIDEFAQGW GPITHVVPN. .ISDQRPYCW HYAPRPCGIV
 HPCUNKCD   SGCPERLATC RPIDTFAQGW GPITYTEPH. .DLDQRPYCW HYAPQPCGIV
  MKC1A     SGCPERMAGC RPIDEFAQGW GPITHVVPN. .ISDQRPYCW HYAPRPCGIV
  NDM59     SGCPERLSAC RGIQAFRIGW GTLRYEDNVT NPEDMRPYCW HYPPKQCGIV
   NZLI     TGCPQRLSSC KPITFFRQGW GPLTDANITG .PSDDRPYCW HYAPRPCDIV
   SA13     TGCPQRMASC RPLAAFDQGW GTISYAAVSG .PSDDKPYCW HYPPRPCGIV
  Th580     TGCSERLGAC KPLEHFQQGW GPITHKSNIT GPSEDRPYCW HYAPRECSVV
Type_3a_CB  TGCPQRLSSC KPITFFKQGW GPLTDANISG .PSDDKPYCW HYAPRPCKVV
 TypeV_D    TGCPQRLSSC KPITFFRQGW GSLTDANVTG .ASADKPYCW HYAPRPCDVV
  VN004     TGCPERMSKC RPLHSFEQGW GPISYVN.IS GSSEDKPYCW HYAPRPCGIV
  VN235     TGCPERLSAC KSITQFAQGW GPVTYAN.VS GSSEDRPYCW HYAPRPCGVV
  VN405     SGCPERLSSC RPLHAFEQGW GPLTYAN.IS GPSNDKPYCW HYPPRPCDIV

            501                                                 550
   BEBE1    PARTVCGPVY CFTPSPVVVG TTDRAGAPTY NWGENETDVF LLNSTRPPKG
   D89815   PASQVCGPVY CFTPSPVVVG TTDRFGAPTY NWGDNETDVL LLNNTRPPHG
ED43type_4  PASSVCGPVY CFTPSPVVVG TTDHVGVPTY TWGENETDVF LLNSTRPPHG
   HC_C2    PASEVCGPVY CFTPSPVVVG TTDRLGVPTY NWGENETDVL LLNNTRPPQG
   HC_G9    PAKKVCGPVY CFTPSPVVVG TTDRAGVPTY RWGANETDVL LLNNSRPPMG
 HCU16326   PTLQVCGPVY CFTPSPVAVG TTDRFGAPTY RWGANETDVL LLNNAGPPQG
 HCV_H_CMR  PAKSVCGPVY CFTPSPVVVG TTDRSGAPTY SWGANDTDVF VLNNTRPPLG
   HCV_J1   PAKSVCGPVY CFTPSPVVVG TTDRSGAPTY NWGANDTDVF VLNNTRPPLG
 HCV_J483   PASQVCGPVY CFTPSPVVVG TTDRSGVPTY SWGENETDVM LLNNTRPPQG
   HCV_J8   PARTVCGPVY CFTPSPVVVG TTDKQGVPTY TWGENETDVF LLNSTRPPRG
  HCV_JK1   PALQVCGPVY CFTPSPVVVG TTDRFGVPTY NWGDNETDVL LLNNTRPPQG
   HCV_JS   PALEVCGPVY CFTPSPVVVG TTDRFGVPTY SWGENETDVL LLNNTRPPQG
 HCV_K1_R1  PAQQVCGPVY CFTPSPVVVG TTDRFGAPTY NWGANETDVL LLNNTRPPQG
 HCV_K1_R2  PASEVCGPVY CFTPSPVVVG TTDRSGAPTY SWGENETDVL LLNNTRAPQG
 HCV_K1_R3  PASEVCGPVY CFTPSPVVVG TTDRSGAPTY NWGGNETDVL LLNNTRPPQG
 HCV_K1_S1  PASQVCGPVY CFTPSPVVVG TTDRFGAPTY NWGANETDVL LLNNTRPPQG
 HCV_K1_S2  PASEVCGPVY CFTPSPVVVG TTDRSGAPTY SWGENETDVL LLNNTRAPQG
 HCV_K1_S3  PASEVCGPVY CFTPSPVVVG TTDRSGAPTY NWGGNETDVL LLNNTRPPQG
   HCV_L2   PASQVCGPVY CFTPSPVVVG TTDRSGAPTY SWGENETDVL LLNNTRPPQG
    HCV_N   PRSEACGPVY CSTPSPVVVG TTDRFGAPTY NWGDNETDVL LLNNTRPPQG
 HCV12083   PARSVCGPVY CFTPSPVVVG TTDKRGNPTY TWGENETDVF MLESLRPPTG
 HCV1480    PARDVCGPVY CFTPSPVVVG TTDRRGCPTY NWGSNETDIL LLNNIRPPAG
 HCVPOLYP   PASEVCGPVY CFTPSPVVVG TTDRHGVPTY SWGENGTDVL LLNNTRPPQG
   HD_1     PASQVCGPVY CFTPSPVVVG TTDRSGVPTY SWGENETDVL LLNNTRPPQG
 HPCCGAA    PAKSVCGPVY CFTPSPVVVG TTDRSGAPTY SWGANDTDVF VLNNTRPPLG
  HPCFG     KPLNVCGPVY CFTPSPVVVG TTDIKGLPTY RFGVNESDVF LLTSLRPPQG
```

```
HPCGENANTI  PASQVCGPVY CFTPSPVVVG TTDRFGAPTY SWGENETDVL ILNNTRPPQG
  HPCGENOM  PASEVCGPVY CFTPSPVVVG TTDRFGVPTY SWGENETDVL LLNNTRPPQG
  HPCHUMR   PASEVCGPVY CFTPSPVVVG TTDRFGVPTY RWGENETDVL LLNNTRPPQG
     HPCJ   PASEVCGPVY CFTPSPVVVG TTDRFGVPTY SWGENETDVL VLNNTRPPQG
    HPCJCG  PASQVCGPVY CFTPSPVVVG TTDRFGAPTY SWGENETDVL LLSNTRPPQG
  HPCJK046  SALNVCGPVY CFTPSPVVLG TTDRRGNPTY TWGANETDVF MMSSLRPPAG
  HPCJK049  PAQSVCGPVY CFTPSPVVVG ATDKRGAPTY TWGENESDVF LLESARPPTE
    HPCJTA  PASQVCGPVY CFTPSPVVVG TTDRSGAPTY NWGANETDVL LLNNTRPPQG
    HPCJTB  PASQVCGPVY CFTPSPVVVG TTDRSGAPTY NWGANETDVL LLNNTRPPQG
    HPCK3A  PALNVCGPVY CFTPSPVVVG TTDAKGAPTY TWGANKTDVF LLESLRPPSG
 HPCPLYPRE  PAKSVCGPVY CFTPSPVVVG TTDRSGAPTY SWGENDTDVF VLNNTRPPLG
   HPCPOLP  SASSVCGPVY CFTPSPVVVG TTDRLGAPTY TWGENETDVF LLNSTRPPQG
     HPCPP  PASQVCGPVY CFTPSPVVVG TTDRFGAPTY NWGNNETDVL LLNNTRPPQG
  HPCUNKCD  PTLQVCGPVY CFTPSPVAVG TTDRFGAPTY RWGANETDVL LLNNAGPPQG
     MKC1A  PASQVCGPVY CFTPSPVVVG TTDRFGAPTY NWGNNETDVL LLNNTRPPQG
     NDM59  SARSVCGPVY CFTPSPVVVG TTDRLGVPTY TWGENETDVF ILNSTRPPGG
      NZLI  PASSVCGPVY CFTPSPVVVG TTDARGVPTY TWGENEKDVF LLKSQRPPSG
      SA13  PARGVCGPVY CFTPSPVVVG TTDRKGNPTY SWGENETDIF LLNNTRPPTG
     Th580  PASSVCGPVY CFTPSPVVVG TTDRLGNPTY NWGENETDVF MLESLRPPQG
 Type_3a_CB PASGVCGPVY CFTPSPVVVG TTDAKGVPTY TWGANDTDVF LLESLRPPGG
   TypeV_D  PALNVCGPVY CFTPSPVVVG TTDRKGVPTY NWGENESDVF LLESLRPPSG
     VN004  PARNVCGPVY CFTPSPVVVG TTDQRGIPTY TWGENVSDVF LLHSARPPLG
     VN235  SARSVCGPVY CFTPSPVVVG TTDRRGVPTY TWGENESDVF LLESLRPPAG
     VN405  PARSVCGPVY CFTPSPVVVG TTDRKGLPTY TWGANESDVF LLRSTRPPRG

            551                                                 600
      BEBE1 AWFGCTWMNG TGFTKTCGAP PCRIRKDFN. .A..SEDLLC PTDCFRKHPG
      D89815 NWFGCTWMNS TGFTKTCGGP PCNIRG.... VG..NNTLTC PTDCFRKHPD
ED43type_4  AWFGCVWMNS TGFTKTCGAP PCEVN..... TN..NGTWHC PTDCFRKHPE
      HC_C2 NWFGCTWMNT TGFTKTCGGP PCNIGG.... AG..NNTLTC PTDCFRKHPE
      HC_G9 NWFGCTWMNS SGFTKTCGAP ACNIGG.... SG..NNTLLC PTDCFRKHPD
   HCU16326 NWFGCTWMNG TGFTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPG
  HCV_H_CMR NWFGCTWMNS TGFTKVCGAP PCVIGG.... VG..NNTLLC PTDCFRKHPE
     HCV_J1 NWFGCTWMNS TGFTKVCGAP PCVIGG.... GG..NNTLHC PTDCFRKHPE
   HCV_J483 NWFGCTWMNS TGFTKTCGGP PCNIGG.... VG..NHTLTC PTDCFRKHPE
     HCV_J8 AWFGCTWMNG TGFTKTCGAP PCRIRKDYN. .S..TIDLLC PTDCFRKHPD
    HCV_JK1 NWFGCTWMNS TGFTKTCGGP PCNIGG.... AG..NNTLTC PTDCFRKHPE
     HCV_JS NWFGCTWMNS TGYTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPE
  HCV_K1_R1 NWFGCTWMNG TGFTKTCGGP PCNIGG.... AG..NNTLTC PTDCFRKHPE
  HCV_K1_R2 NWFGCTWMNG TGFTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPE
  HCV_K1_R3 NWFGCTWMNS TGFTKTCGGP PCNIGG.... AG..NNTFTC PTDCFRKHPE
  HCV_K1_S1 NWFGCTWMNG TGFTKTCGGP PCNIGG.... AG..NDTLTC PTDCFRKHPE
  HCV_K1_S2 NWFGCTWMNG TGFTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPE
  HCV_K1_S3 NWFGCTWMNS TGFTKTCGGP PCNIGG.... AG..NNTLTC PTDCFRKHPE
     HCV_L2 NWFGCTWMSS TGFTKTCGGP PCNIGG.... AG..NNTLTC PTDCFRKHPE
      HCV_N NWFGCTWMNS TGFTKTCGAP PCNIGG.... VG..NNTLTC PTDCFRKHPE
   HCV12083 GWFGCTWMNS TGFTKTCGAP PCQIVPGNYN SS..ANELLC PTDCFRKHPE
    HCV1480 NWFGCTWMNS TGFVKNCGAP PCNLGP.... TG..NNSLKC PTDCFRKHPD
   HCVPOLYP NWFGCTWMNG TGFTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPE
       HD_1 NWFGCTWMNS TGFTKTCGGP PCGIGG.... VG..NNTLIC PTDCFRKHPE
    HPCCGAA NWFGCTWMNS TGFTKVCGAP PCVIGG.... VG..NNTLLC PTDCFRKYPE
      HPCFG RWFGCVWMNS TGFVKTCGAP PCNIYGGMKD IEANQTHLKC PTDCFRKHHD
HPCGENANTI NWFGCTWMNS TGFTKTCGGP PCNIGG.... GG..NNTLVC PTDCFRKHPE
  HPCGENOM NWFGCTWMNG TGFTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPE
   HPCHUMR NWFGCTWMNS TGFTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPE
     HPCJ  NWFGCTWMNG TGFTKTCGGP PCHIGG.... RG..NNTLTC PTDCFRKHPE
    HPCJCG NWFGCTWMNS TGFTKTCGGP PCNIGG.... VG..NNTLVC PTDCFRKHPE
  HPCJK046 GWYGCTWMNT SGFVKTCGAP PCNIRPNPEE NR..TETLRC PTDCFRKHPG
  HPCJK049 PWFGCTWMNG SGYVKTCGAP PCHIYGGREG KS..NNSLVC PTDCFRKHPD
    HPCJTA NWFGCTWMNS TGFTKTCGGP PCNIGG.... VG..NLTLTC PTDCFRKHPE
    HPCJTB NWFGCTWMNS TGFTKTCGGP PCNIGG.... VG..NLTLTC PTDCFRKHPE
```

```
HPCK3A      RWFGCTWMNS TGFVKTCGAP PCNIYGDGRD AQ.NESDLFC PTDCFRKHPE
HPCPLYPRE   NWFGCTWMNS TGFTKVCGAP PCVIGG.... AG..NNTLHC PTDCFRKHPD
HPCPOLP     SWFGCTWMNS TGYTKTCGAP PCRIRADFN. .A..SMDLLC PTDCFRKHPD
HPCPP       NWFGCTWMNG TGFTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPE
HPCUNKCD    NWFGCTWMNG TGFTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPG
MKC1A       NWFGCTWMNG TGFTKTCGGP PCNIGG.... VG..NNTLTC PTDCFRKHPE
NDM59       SWFGCTWMNS TGFTKTCGAP PCRIRADFN. .A..SMDLLC PTDCFRKHPD
NZLI        RWFGCSWMNS TGFLKTCGAP PCNIYGGEGN PH.NESDLFC PTDCFRKHPE
SA13        NWFGCTWMNS TGFVKTCGAP PCNLGP.... TG..NNSLKC PTDCFRKHPD
Th580       GWFGCTWMNS TGFTKTCGAP PCQLIPGDYN SS..SNQLLC PTDCFRKHPE
Type_3a_CB  RWFGCTWMNS TGFVKTCGAS PCDIYGGGGN SG.NESDLFC PTDCFRKHPE
TypeV_D     RWFGCAWMNS TGFLKTCGAP PCNIYGGGGN PN.NESHLFC PTDCFRKHPD
VN004       AWFGCTWMNS SGFVKTCGAP PCRIKPTIN. ....ETDLVC PTDCFRKHPD
VN235       AWYGCTWMNS TGYTKTCGAP PCHIGP.PDQ ........FC PTDCFRKHPE
VN405       SWFGCTWMNS TGFVKTCGAP PCNTRPVGSG ....NDTLVC PTDCFRKHPE

            601                                                  650
BEBE1       ATYIKCGAGP WLTPRCLVDY PYRLWHYPCT VNYTIYKVRM FVGGIEHRLQ
D89815      ATYTKCGSGP WLTPRCLVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLD
ED43type_4  TTYAKCGSGP WITPRCLIDY PYRLWHFPCT ANFSVFNIRT FVGGIEHRMQ
HC_C2       ATYTKCGSGP WLTPRCLVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLD
HC_G9       ATYSRCGSGP WLTPRCLVDY PYRLWHYPCT VNYTIFKIRM FVGGVEHRLD
HCU16326    ATYTKCGSGP WLTPRCLVDY PYRLWHYPCT VNFTIFKVRM YVGGAEHRLD
HCV_H_CMR   ATYSRCGSGP WITPRCMVDY PYRLWHYPCT INYTIFKVRM YVGGVEHRLE
HCV_J1      ATYSRCGSGP WITPRCLVDY PYRLWHYPCT INYTIFKVRM YVGGVEHRLD
HCV_J483    ATYTKCGSGP WLTPRCLVDY PYRLWHYPCT FNFSIFKVRM YVGGVEHRLN
HCV_J8      ATYLKCGAGP WLTPRCLVDY PYRLWHYPCT VNFTIFKARM YVGGVEHRFS
HCV_JK1     ATYTKCGSGP WLTPRCMVDY PYRLWHYPCT FNFTIFKIRM YVGGVEHRLN
HCV_JS      ATYTKCGSGP WLTPRCLVHY PYRLWHYPCT VNFTIFKVRM YVGGIEHRLD
HCV_K1_R1   ATYAKCGSGP WLTPRCIVDY PYRLWHYPCT VNFSIFKVRM YVGGVEHRLT
HCV_K1_R2   ATYTKCGSGP WLTPRCMVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLN
HCV_K1_R3   ATYTKCGSGP WLTPRCIVDY PYRLWYYPCT VNFTIFTIRM YVGGVEHRLK
HCV_K1_S1   ATYAKCGSGP WLTPRCIVDY PYRLWHYPCT VNFSIFKVRM YVGGVEHRLT
HCV_K1_S2   ATYTKCGSGP WLTPRCMVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLN
HCV_K1_S3   ATYTKCGSGP WLTPRCIVDY PYRLWYYPCT VNFTIFTIRM YVGGVEHRLK
HCV_L2      ATYTKCGSGP WLTPRCLVDY PYRLWHYPCT VNFTTFKVRM YVGGVEHRLI
HCV_N       ATYSKCGSGP WLTPRCMVDY PYRLWHYPCT VNFSIFKVRM YVGGVEHRLN
HCV12083    ATYQRCGSGP WVTPRCLVDY AYRLWHYPCT VNFTLHKVRM FVGGTEHRFD
HCV1480     ATYTRCGSGP WLTPRCLVHY PYRLWHYPCT VNYTIFKVRM FIGGLEHRLE
HCVPOLYP    ATYTKCGSGP WLTPRCMVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLS
HD_1        ATYTRCGSGP WLTPRCMVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLN
HPCCGAA     ATYSRCGSGP RITPRCMVDY PYRLWHYPCT INYTIFKVRM YVGGVEHRLE
HPCFG       ATFTRCGSGP WLTPRCLVDY PYRLWHYPCT VNFSIFKVRM FVGGHEHRFS
HPCGENANTI  ATYTKCGSGP WLTPRCMVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLN
HPCGENOM    ATYTRCGSGP WLTPRCLVDY PYRLWHYPCT VNFAIFKVRM YVGGVEHRLD
HPCHUMR     ATYTKCGSGP WLTPRCMVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLN
HPCJ        ATYTKCGSGP WLTPRCMVDY PYRLWHYPCT VNFTTFKVRM YVGGVEHRLI
HPCJCG      ATYTKCGSGP WLTPRCMVDY PYRLWHYPCT VNFTVFKVRM YVGGVEHRLN
HPCJK046    ATYAKCGSGP WLTPRCLVDY PYRLWHYPCT VNYTLHKVRM YIAGSEHRFT
HPCJK049    ATYNRCGAGP WLTPRCLVDY PYRLWHYPCT VNYTIFKVRM FVGGLEHRFN
HPCJTA      ATYTKCGSGP WLTPRCIVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLS
HPCJTB      ATYTKCGSGP WLTPRCIVDY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLS
HPCK3A      ATYSRCGAGP WLTPRCLVDY PYRLWHYPCT VNFTLFKVRM FVGGFEHRFT
HPCPLYPRE   ATYSRCGSGP WITPRCLVDY PYRLWHYPCT INYTIFKIRM YVGGVEHRLE
HPCPOLP     TTYIKCGSGP WLTPRCLIDY PYRLWHYPCT VNYTIFKIRM YVGGVEHRLT
HPCPP       ATYTKCGSGP WLTPRCLVHY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLE
HPCUNKCD    ATYTKCGSGP WLTPRCLVDY PYRLWHYPCT VNFTIFKVRM YVGGAEHRLD
MKC1A       ATYTKCGSGP WLTPRCLVHY PYRLWHYPCT VNFTIFKVRM YVGGVEHRLE
NDM59       ATYIKCGSGP WLTPRCLVDY PYRLWHYPCT INYTIFKIRM YVGGVEHRLT
NZLI        TTYSRCGAGP WLTPRCMVDY PYRLWHYPCT VDFRLFKVRM FVGGFEHRFT
SA13        ATYTKCGSGP WLTPRCLVHY PYRLWHYPCT LNYTIFKVRM YIGGLEHRLE
```

```
     Th580  ATYQKCGSGP WLTPRCLVDY PYRLWHYPCT VNYTIHKVRM FIGGVEHRFD
Type_3a_CB  ATYSRCGAGP WLTPRCMVDY PYRLWHYPCT VNFTLFKVRM FVGGFEHRFT
   TypeV_D  ATYSRCGAGP WLTPRCMVDY PYRLWHYPCT VNFTLFKVRM FVGGFEHRFT
     VN004  ASFVKCGSGP WLTPRCMVDY PYRLWHYPCT VNFTIHKVRV FVGGVEHRFN
     VN235  ATYRKCGSGP WLTPRCLVDY PYRLWHYPCT VNYTIHKVRL FINGLEHRFD
     VN405  ATYARCGSGP WLTPRCLVNY PYRLWHYPCT VNYTIHKVRM FVGGIEHRFE

            651                                                  700
     BEBE1  AACNFTRGDR CNLEDRDRSQ LSPLLHSTTE WAILPCSYTD LPALSTGLLH
     D89815 AACNWTRGER CDLEDRDRAE LSPLLLSTTE WQILPCSYTT LPALSTGLIH
ED43type_4  AACNWTRGEV CGLEHRDRVE LSPLLLTTTA WQILPCSFTT LPALSTGLIH
     HC_C2  AACNWTRGER CALEDRDRSE LSPLLLSTTE WQILPCSFTT LPALSTGLIH
     HC_G9  AACNWTRGER CDLDDRDRAE LSPLLLSTTQ WQVLPCSFTT LPALSTGLIH
  HCU16326  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQVLPCSFTT LPALSTGLIH
 HCV_H_CMR  AACNWTRGER CDLEDRDRSE LSPLLLSTTQ WQVLPCSFTT LPALSTGLIH
     HCV_J1  AACNWTRGER CDLEDRDRSE LSPLLLSTTQ WQVLPCSFTT LPALSTGLIH
  HCV_J483  AACNWTRGER CNLEDRDRSE LSPLLLSTTE WQILPCAFTT LPALSTGLIH
     HCV_J8  AACNFTRGDR CRLEDRDRGQ QSPLLHSTTE WAVLPCSFSD LPALSTGLLH
    HCV_JK1  AACNWTRGER CNIEDRDRSE LSPLLLSTTE WQILPCSFTT LPALSTGLIH
     HCV_JS  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQILPCSFTT LPALSTGLIH
  HCV_K1_R1  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQILPCSFTT LPALSTGLIH
  HCV_K1_R2  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQILPCSFTT LPVLSTGLIH
  HCV_K1_R3  AACNWTRGER CNLEDRDRAE LSPLLLSTTE WQILPCSFTT LPALSTGLIH
  HCV_K1_S1  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQILPCSFTT LPALSTGLIH
  HCV_K1_S2  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQILPCSFTT LPVLSTGLIH
  HCV_K1_S3  AACNWTRGER CNLEDRDRAE LSPLLLSTTE WQILPCSFTT LPALSTGLIH
     HCV_L2  AACNWTRGER CNLEDRDRSE LSPLLLSTTE WQILPCSYTT LPALSTGLIH
      HCV_N  AACNWTRGER CDLDDRDRSE LSPLLLSTTE WQVLPCSFTT LPALSTGLIH
  HCV12083  VACNWTRGER CELHDRNRIE MSPLLFSTTQ LSILPCSFST MPALSTGLIH
   HCV1480  AACNWTYGER CDLEDRDRAE LSPLLHTTTQ WAILPCSFTP TPALSTGLIH
  HCVPOLYP  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQVLPCSFTT LPALSTGLIH
      HD_1  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQVLPCSFTT LPALSTGLIH
   HPCCGAA  AACNWTRGER CDLEDRDRSE LSPLLLSTTQ WQVLPCSFTT LPALSTGLIH
     HPCFG  AACNWTRGER CDLEDRDRSE QQPLLHSTTD SLILPCSFTP MRRLSTGLIH
HPCGENANTI  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQILPCSFTG LPALSTGLIH
  HPCGENOM  AACNWTRGER CNLEDRDRSE LSPLLLSTTE WQILPCAFTT LPALSTGLIH
   HPCHUMR  AACNWTRGER CDLEDRDRPE LSPLLLSTTE WQVLPCSFTT LPALSTGLIH
      HPCJ  AACNWTRGDR CNLEDRDRSE LSPLLLSTTE WQILPCSFTT LPALSTGLIH
    HPCJCG  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQILPCSFTT LPALSTGLIH
  HPCJK046  AACNWTRGER CDLADRDRIE MSPLLFSTTE LAILPCSFTT MPALSTGLIH
  HPCJK049  AACNWTRGER CNLEDRDRSE MYPLLHSTTE QAILPCSFVP IPALSTGLIH
    HPCJTA  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQTLPCSFTT LPALSTGLIH
    HPCJTB  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQTLPCSFTT LPALSTGLIH
    HPCK3A  AACNWTRGER CDIEDRDRSE QHPLLHSTTE LAILPCSFTP MPALSTGLIH
 HPCPLYPRE  AACNWTRGER CDLEDRDRSE LSPLLLTTTQ WQVLPCSFTT LPALSTGLIH
   HPCPOLP  AACNFTRGDR CNLEDRDRSQ LSPLLHSTTE WAILPCTYSD LPALSTGLLH
     HPCPP  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQVLPCSFTT LPALSTGLIH
  HPCUNKCD  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQVLPCSFTT LPALSTGLIH
     MKC1A  AACNWTRGER CDLEDRDRSE LSPLLLSTTE WQVLPCSFTT LPALSTGLIH
     NDM59  AACNFTRGDP CNLEDRDRSQ LSPLLHSTTE WAILPCSYSD LPALSTGLLH
      NZLI  AACNWTRGER CDIEDRDRSE QHPLLHSTTE LAILPCSFTP MPALSTGLIH
      SA13  VACNWTRGER CDLEDRDRAE LSPLLHTTTQ WAILPCSFTP TPALSTGLIH
     Th580  AACNWTRGDR CDLYDRDRIE MSPLLFSTTQ LAILPCSFTT MPALSTGLIH
Type_3a_CB  AACNWTRGER CDIEDRDRSE QHPLLHSTTE LAILPCSFTP MPALSTGLIH
   TypeV_D  AACNWTRGER CNIEDRDRSE QHPLLHSTTE LAILPCSFTP MPALSTGLIH
     VN004  AACNWTRGDR CELDDRDRFE MSPLLFSTTQ LAILPCSFTT MPALSTGLIH
     VN235  AACNWTRGER CELEDRDRIE MSPLLFSTTE LAILPCSFTT MPALSTGLVH
     VN405  AACNWTRGER CELDDRDRVE MSPLLFSTTQ LSILPCSFTT MPALSTGLIH

            701                                                  750
     BEBE1  LHQNIVDVQY LYGLSPAITK YVVKWEWVVL LFLLLADARV CACLWMLLLL
```

```
D89815          LHQNIVDIQY LYGIGSAVVS IAIKWEYVVL LFLLLADARV CACLWMMLLI
ED43type_4      LHQNIVDVQY LYGVGSAVVS WALKWEYVVL AFLLLADARV SAYLWMMFMV
HC_C2           LHRNIVDVQY LYGIGSAVVS FAIKWEYVLL LFLLLADARV CACLWMMLLI
HC_G9           LHQNIVDVQY LYGLSSAVTS WVIKWEYVVL LFLLLADARI CACLWMMLLI
HCU16326        LHQNIVDIQY LYGIGSAVVS FAIKWEYIVL LFLLLADARV CACLWMMLLV
HCV_H_CMR       LHQNIVDVQY LYGVGSSIAS WAIKWEYVVL LFLLLADARV CSCLWMMLLI
HCV_J1          LHQNIVDVQY LYGVGSSIAS WAIKWEYVVL LFLLLADARV CSCLWMMLLI
HCV_J483        LHQNIVDVQY LYGVGSAFVS FAIKWEYILL LFLLLADARV CACLWMMLLI
HCV_J8          LHQNIVDVQY LYGLSPALTR YIVKWEWVIL LFLLLADARI CACLWMLIIL
HCV_JK1         LHQNIVDVQY LYGVGSAVVS IVIKWEYVLL LFLLLADARV CACLWMMLLI
HCV_JS          LHRNIVDVQY LYGIGSAVVS FAIKWEYVLL LFLLLADARV CACLWMMLLI
HCV_K1_R1       LHQNIVDVQY LYGVGSAVVS FAIKWEYVLL LFLLLADARV CACLWMMLLV
HCV_K1_R2       LHQNIVDVQY LYGVGSAVVS FVIKWEYILL LFLLLADARV CACLWMMLLI
HCV_K1_R3       LHQNIVDVQY LYGVGSAVVS VVIRWEYVLL LFLLLADARV CACLWMMLLI
HCV_K1_S1       LHQNIVDVQY LYGVGSAVVS FAIKWEYVLL LFLLLADARV CACLWMMLLV
HCV_K1_S2       LHQNIVDVQY LYGVGSAVVS FVIKWEYILL LFLLLADARV CACLWMMLLI
HCV_K1_S3       LHQNIVDVQY LYGVGSAVVS VVIRWEYVLL LFLLLADARV CACLWMMLLI
HCV_L2          LHQNIVDVQY LYGIGSAVVS FVIKWEYVLL FFLLLADARV CACLWMILLI
HCV_N           LHQNIVDVQY LYGIGSAVVS FAIKWEYVVL LFLLLADARV CACLWMMLLI
HCV12083        LHQNIVDVQY LYGVSTNVTS WVVKWEYIVL MFLVLADARI CTCLWLMLLI
HCV1480         LHQNIVDTQY LYGLSSSIVS WAVKWEYIML VFLLLADARI CTCLLILLLI
HCVPOLYP        LHQNIVDVQY LYGIGSVVVS FAIKWEYVVL LFLLLADARV CACLWMMLLI
HD_1            LHQNVVDVQY LYGIGSVVVS FAIKWEYILL LFLLLADARV CACLWMMLLI
HPCCGAA         LHQNIVDVQY LYGVGSSIAS WAIKWEYVVL LFLLLADARV CSCLWMMLLI
HPCFG           LHQNIVDVQY LYGVGSAVVG WALKWEFVVL VFLLLADARV CVALWMMLLI
HPCGENANTI      LHQNVVDVQY LYGIGSAVVS FAIKWEYILL LFLLLADARV CACLWMMLLI
HPCGENOM        LHQNIVDVQY LYGIGSAVAS FAIKWEYVLL LFLLLADARV CACLWMMLLI
HPCHUMR         LHQNIVDVQY LYGIGSAVVS FAIKWEYVLL LFLLLADARV CACLWMMLLI
HPCJ            LHQNIVDVQY LYGIGSAVVS IAIRWEYVLL LFLLLADARV CACLWMMLLI
HPCJCG          LHRNIVDVQY LYGIGSAVVS FAIKWEYILL LFLLLADARV CACLWMMLLI
HPCJK046        LHQNVVDVQY LYGLSTSIVN WAIKWEYVVL LFLVLADSRI CLALWLMLLI
HPCJK049        LHQNIVDVQY LYGISSGLVG WAIKWEFVIL IFLLLADARV CVVLWMMMLI
HPCJTA          LHQNIVDVQY LYGIGSAVVS FVIKWEYIVL LFLLLADARV CACLWMMLLI
HPCJTB          LHQNIVDVQY LYGIGSAVVS FVIKWEYIVL LFLLLADARV CACLWMMLLI
HPCK3A          LHQNIVDVQY LYGIGSGMVG WALKWEFVIL IFLLLADARV CVALWLILTI
HPCPLYPRE       LHQNIVDVQY LYGVGSSIAS WAIKWEYVVL LFLLLADARV CSCLWMMLLI
HPCPOLP         LHQNIVDVQF MYGLSPALTK YIVRWEWVVL LFLLLADARV CACLWMLILL
HPCPP           LHQNIVDVQY LYGIGSAVVS FAIKWEYILL LFLPLADARV CAWLWMMLLI
HPCUNKCD        LHQNIVDIQY LYGIGSAVVS FAIKWEYIVL LFLLLADARV CACLWMMLLV
MKC1A           LHQNIVDVQY LYGIGSAVVS FAIKWEYILL LFLPLADARV CAWLWMMLLI
NDM59           LHQNIVDVQY MYGLSPALTK YIVRWEWVVL LFLLLADARV CACLWMLILL
NZLI            LHQNIVDVQY LYGVGSGMVG WALKWEFVIL VFLLLADARV CVALWLMLMI
SA13            LHQNIVDTQY LYGLSSSIVS WAVKWEYIVL AFLLLADARI CTCLWIMLLV
Th580           LHQNIVDVQY LYGVSSSIVS WAVKWEYVVL MFLVLADARI CTCLWLMLLV
Type_3a_CB      LHQNIVDVQY LYGVGSGMVG WALKWEFVIL IFLLLADRRV CVALWLMLMI
TypeV_D         LHQNIVDVQY LYGVGSGMVG WALKWEFVIL IFLLLADARV CVALWLMLMI
VN004           LHQNIVDIQY LYGVSTAVVS WAMKWEYVVL AFLVLADARV CACLWLMFLV
VN235           LHQNIVDIQY LYGLAPALVS WAVRWEYVVL AFLLLADARI CACLWMVLLI
VN405           LHQNIVDVQY LYGVSSAVVS WAVKWEYIVL AFLVLAVARV CACLWLMFLV

                751                                                 800
BEBE1           GQAEAALEKL VILHAASAAS SNGLLYFILF FVAAWCIKGR AVPMVTYTLL
D89815          AQAEAALENL VVLNAASVVG AHGMLPFFMF FCAAWYMKGR LVPGAAYAFY
ED43type_4      SQVEAALSNL ININAASAAG AQGFWYAILF ICIVWHVKGR FPAAAAYAAC
HC_C2           AQAEAALENL VVLNAASVAG VHGILSFLVF FCAAWYIKGR LVPGAAYAFY
HC_G9           SQVEAALENL IVLNAASLVG THGIVPFFIF FCAAWYLKGK WAPGLAYSVY
HCU16326        AQAEAALENL VVLNAASVAG AHGILSFIVF FCAAWYIKGR LVPGAAYALY
HCV_H_CMR       SQAEAALENL VILNAASLAG THGLVSFLVF FCFAWYLKGR WVPGAVYAFY
HCV_J1          SQAEAALENL VILNAASLAG TRGLVSFLVF FCFAWYLKGR WVPGAAYALY
HCV_J483        AQAEAALENL VVLNAASVAG AHGILSFLVF FCAAWYIKGR LAPGAAYAFY
HCV_J8          GQAEAALEKL IILHSASAAS ANGPLWFFIF FTAAWYLKGR VVPVATYSVL
```

```
HCV_JK1     AQAEAALENL VVLNAASVAG AHGILSFLVF FCAAWYIKGR LVPGAAYAFY
HCV_JS      AQAEAALENL VVLNAASVAG AHGILSFLVF FCAAWCIKGR LVPGAAYAFY
HCV_K1_R1   AQAEAALENL VVLNAASLAG VHGILSFLVF FCAAWYIKGK LVPGAAYALY
HCV_K1_R2   AQAEAALENL VVLNAASVAG AHGFLSFLVF FCAAWYIKGK LVPGAAYAFY
HCV_K1_R3   AQAEAALENL VILNAASVAG AHGILSFLVF FCAAWYIKGK LVPGAAYAFY
HCV_K1_S1   AQAEAALENL VVLNAASLAG VHGILSFLVF FCAAWYIKGK LVPGAAYAFY
HCV_K1_S2   AQAEAALENL VVLNAASVAG AHGFLSFLVF FCAAWYIKGK LVPGAAYALY
HCV_K1_S3   AQAEAALENL VILNAASVAG AHGILSFLVF FCAAWYIKGR LVPGAAYAFY
HCV_L2      AQAEAALENL VVLNAASVAG AHGILSFLVF FCAAWYIKGR LVPGAAYASY
HCV_N       ARAEAALENL VALNAASVAG AHGILSFLVF FCAAWYIKGR LVPGAAYAFY
HCV12083    STVEAAVERL VVLNAASAAG TAGWWWAVLF LCCVWYVKGR LVPACTYMAL
HCV1480     CQAEATCKNV IVLNAAAAAG NHGFFWGLLV VCLAWHVKGR LVPGATYLCL
HCVPOLYP    AQAEAALENL VVLNAASVAG AHGILSFLVF FCAAWYIKGR LVPGAAYAIY
HD_1        AQAEAALENL VVLNAASLAG THGILSFLVF FCAAWYIKGR LVPGAAYALY
HPCCGAA     SQAEAALENL VILNAASLAG THGLVSFLVF FCFAWYLKGR WVPGAVYALY
HPCFG       SQAEAAMENL VMLNALSAAG QQGYVWYLVA FCAAWHIRGK LVPLITYGLT
HPCGENANTI  AQAEAALENL VVFNAASVAG MHGTLSFLVF FCAAWYIKGR LVPGAAYALY
HPCGENOM    AQAEAALENL VVLNAASVAD AHGILSFLVF FCAAWYIKGR LVPGAAYAFY
HPCHUMR     AQAEAALENL VVLNSASVAG AHGILSFLVF FCAAWYIKGR LVPGATYALY
HPCJ        AQAEAALENL VVLNAASVAG AHGILSFLVF FCAAWYIKGR LVPGAAYALY
HPCJCG      AQAEATLENL VVLNAASVAG AHGLLSFLVF FCAAWYIKGR LVPGAAYALY
HPCJK046    GQAEAALENL IVLNAASAAA TRGWECFLLF MCWAWYVRGR VVPAVTYGLL
HPCJK049    SQAEAALENL IVLNAISAAG THGIWWSLVA FCVAWHVRGR IFPIAVYSIV
HPCJTA      AQAEAALENL VVLNAASLAG ADGILSFLVF FCAAWYIKGR LVPGAAYALY
HPCJTB      AQAEAALENL VVLNAASLAG AHGILSFLVF FCAAWYIKGR LVPGAAYALY
HPCK3A      SQAEAALENL VTLNAVAAAG THGIGWYLVA FCAAWYVRGK LVPLVTYSLT
HPCPLYPRE   SQAEAALENL VILNAASLAG THGLVSFLVF FCFAWYLKGK WVPGAVYTFY
HPCPOLP     GQAEAALEKL VVLHAASAAS CNGFLYFVIF FVAAWYIKGR VVPLATYSLT
HPCPP       AQAEAALENL VVLNAASMAG AHGILSFLVF FCAAWYIKGR LVPGAAYAFY
HPCUNKCD    AQAEAALENL VVLNAASVAG AHGILSFIVF FCAAWYIKGR LVPGAAYALY
MKC1A       AQAEAALENL VVLNAASMAG AHGILSFLVF FCAAWYIKGR LVPGAAYAFY
NDM59       GQAEAALEKL VVLHAASAAS CNGFLYFVIF FVAAWYIKGR AVPLAAYSLT
NZLI        SQTEAALENL VTLNAVAAAG THGIGWYLVA FCAAWYVRGK LVPLVTYSLT
SA13        CQAEAALENV IVLNAAAAAG THGFFWGLLV ICFAWHFKGR LVPGATYLCL
Th580       GKVEAALERL VVLNAASAAG TAGWCWTLIF LCCVWHVKGR LVPACTYTAL
Type_3a_CB  TQAEAALENL VTLNAVAAAG THGIGWYLVA FCAAWYVRGK LVPVVTYSLT
TypeV_D     SQAEAALENL VTLKAVAAAG THGIGWYLVA FCAAWHVRGK LVPLVTYSLT
VN004       GQAEAALENV IVLNAASAAS CQGLLWGLIF ICCAWHVRGR AVPVTTYALL
VN235       SQVEAALENL IVLNAASAAS SQGWIYCLVF ICCAWYIKGR VVPGATYAIL
VN405       GQAEAALENL IVLNATSAAG SQGWVWGVVF ICAAWYIRGR AAPITTYAIL

            801                                                 850
BEBE1       GCWSFVLLLM ALPHQAYALD AAEQGQIGMA LLIAITAFTI TPAYKILLSR
D89815      GVWPLLLLLL ALPPRAYAMD REMVASCGGG VFVGLALLTL SPYCKVFLAR
ED43type_4  GLWPCFLLLL MLPERAYAYD QEVAGSLGGA IVVMLTILTL SPHYKLWLAR
HC_C2       GVWPLLLLLL ALPPRAYAMA REMAASCGGA VFVGLALLTL SPYYKVFLAR
HC_G9       GMWPLLLLLL ALPQRAYALD QELAASCGAT VFICLAVLTL SPYYKQYMAR
HCU16326    GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFVGLVLLTL SPHYKVFLAR
HCV_H_CMR   GMWPLLLLLL ALPQRAYALD TEVAASCGGV VLVGLMALTL SPYYKRYISW
HCV_J1      GMWPLLLLLL ALPQRAYALD TEVAASCGGV VLVGLMALTL SPYYKRCISW
HCV_J483    GVWPLLLLLL ALPPRAYAMD REMAASCGGA VLVGLVFLTL SPYYKVFLTR
HCV_J8      GLWSFLLLVL ALPQQAYALD AAEQGELGLA ILVIISIFTL TPAYKILLSR
HCV_JK1     GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFVGLALLTL SPHYKAFLAR
HCV_JS      GVWPLLLLLL ALPPRAFAMD REMAASCGGA VFVGLALLTL SPHYKVFLAR
HCV_K1_R1   GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFIGLALLTL SPYYKVFLAK
HCV_K1_R2   GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFVGLALLTL SPHYKVVLAR
HCV_K1_R3   GVWPLLLLLL SLPPRAYAMD REMAASCGGA VFIGLMLLTL SPHYKAFLAK
HCV_K1_S1   GVWPLLLLLL ALPPRAYAMD REMAASCGGV VFVGLALLTL SPYYKVFLAR
HCV_K1_S2   GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFVGLALLTL SPHYKVVLAR
HCV_K1_S3   GVWPLLLLLL SLPPRAYAMD REMAASCGGA VFIGLILLTL SPHYKAFLAK
HCV_L2      GVWPLLLLLL ALPPRAYAMD QGMAASSGGT VLVGLMLLTL SPYYKVVLAR
```

```
          GAWPLLLLLL TLPPRAYAMD REMAASCGGA VFVGLALLTL SPYYKVFLAR
HCV12083  GMWPLLLTIL ALPPRAYAMD NEQAASLGAV GLLVITIFSI TPMYKKLLNC
HCV1480   GVWPLLLVRL LRPHRALALD SSDGGTVGCL VLIVLTIFTL TPGYKKKVVL
HCVPOLYP  GVWPLLLLLL ALPPRAYALD REMAASCGGA VFVGLVLLTL SPHYKEFLAR
HD_1      GVWPLLLLLL ALPPRAYALD REMAASCGGA VFVGLIILTL SPHYKAFLAR
HPCCGAA   GMWPLLLLLL ALPQRAYALD TEVAASCGGV VLVGLMALTL SPYYKRYISW
HPCFG     GLWPLALLDL LLPQRAYAWT GEDDATIGAG VLLLLGFFTL SPWYKHWIGR
HPCGENANTI GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFVGLVLLTL SPHYKMFLAR
HPCGENOM  GVWPLLLLLL TLPPRAYAMD REMAASCGGA VFVGLVFLTL SPYYKVFLAR
HPCHUMR   GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFVGLVLLTL SPYYKVFLAR
HPCJ      GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFMGLAFLTL SPHYKVFLAR
HPCJCG    GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFVGLVLLTL SPYYKVFLAR
HPCJK046  NLWPLLLLVL LLPHRAYAYD GVQAGSIGAA VIAALTIFSL TPAYKTLLAH
HPCJK049  GLWPLLLLVL MLPYRAYAWT GTDTSTLGAG VLSLFALFTL SPWYKHWIAR
HPCJTA    GVWPLLLLLL ALPPRAYAMD REMAASCGGV VFVGLILLTL SPHYKVFLAR
HPCJTB    GVWPLLLLLL ALPPRAYAMD REMAASCGGV VFLGLILLTL SPHYKVFLAR
HPCK3A    GLWSLALLVL LLPQRAYAWS GEDSATLGAG ILVLFGFFTL SPWYKHWIAR
HPCPLYPRE GMWPLLLLLL ALPQRAYALD TEVAASCGGV VLVGLMALTL SPYYKRYISW
HPCPOLP   GLWSFGLLLL ALPQQAYAYD ASVHGQIGAA LLVLITLFTL TPGYKTLLSR
HPCPP     GVWPLLLLLM ALPARAYAMD REMAASCGGA VFVGLVLLTL SPYYKVFLAK
HPCUNKCD  GVWPLLLLLL ALPPRAYAMD REMAASCGGA VFVGLVLLTL SPHYKVFLAR
MKC1A     GVWPLLLLLM ALPARAYAMD REMAASCGGA VFVGLVLLTL SPYYKVFLAK
NDM59     GLWPFCLLLL ALPQQAYAYD ASVHGQTGVA LLILITLFTL TPGYKILLSR
NZLI      GLWSLALLVL LLPQRAYAWS GEDSATLGAG VLVLFGFFTL SPWYKHWIGR
SA13      GIWPLLLLLF LLPQRALALD SSDGGTVGCL VLTILTIFTL TPGYKKMVVL
Th580     GMWPILLVIL ALPQRAYAWD NSQAASLGVV ALLVLTIFTL SPMYKQLLTH
Type_3a_CB GLWPLALLVL MLPQRAYAWS GEDSATLGAG VLALFGFFTL SPWYKHWIGR
TypeV_D   GLWSLALLVL LLPQRAYAWS GEDSATLGAG VLVLFGFFTL SPWYKHWIGR
VN004     QLWPLLLLIL ALPRRAYAFD SEQAASAGLL VLGLITIFTL TPAYKQLLIS
VN235     HLWPLLLLVL ALPQRAYAQD REQGASIGVV VIAAITIFTL TPAYKTMLVH
VN405     QLWPLLLLVL ALPRRAYAYN GEEAASLGML AIVIITIFTL TPAYKTLLIS

          851                                                 900
BEBE1     CLWWTCYMLV LAEALIQDWI PPLQARGGRD GVIWAMTMFY PGVVFDITKW
D89815    LIWWLQYFIT KAEAHLQVSL PPLNVRGGRD AIILLMCAVH PELIFDITKL
ED43type_4 GLWWIQYFIA RTEAVLHVYI PSFNVRGPRD SVIVLAVLVC PDLVFDITKY
HC_C2     LIWWLQYFIT RAEAHLQVWI PPLNVRGGRD AIILLTCAAH PELIFDITKL
HC_G9     GIWWLQYMLT RAEALLQVWV PPLNARGGRD GVVLLTCVLH PHLLFEITKI
HCU16326  FIWWLQYLIT RTEAHLQVWV PPLNVRGGRD AIILLTCVVH PELIFDITKY
HCV_H_CMR CMWWLQYFLT RVEAQLHVWV PPLNVRGGRD AVILLMCVVH PTLVFDITKL
HCV_J1    CLWWLQYFLT RVEAQLHVWV PPLNVRGGRD AVILLMCVVH PTLVFDITKL
HCV_J483  LIWWLQYFIT RAEAHMQVWV PPLNVRGGRD AIILLTCAVH PELIFDITKL
HCV_J8    SVWWLSYMLV LAEAQIQQWV PPLEVRGGRD GIIWVAVILH PRLVFEVTKW
HCV_JK1   LIWWLQYFIT RVEACLQVWV PPLNVRGGRD AIILLTCAVH SELIFDITKI
HCV_JS    LMWRLQYLIT RAEAHLQIWV PPLNIRGGRD AIILLMCAVH PELIFDITKL
HCV_K1_R1 LIWWLQYLIT RAEAQLQVWI PPLNVRGGRD AIILLTCAVH PELIFDITKI
HCV_K1_R2 LIWWLQYFIT RAEACLQVWI PPLNVRGGRD AIILLTCALH PELIFSITKI
HCV_K1_R3 LIWWLQYFIT RAEAHLQVWI PPLNVRGGRD AIILLTCAVH SELVFEITKI
HCV_K1_S1 LIWWLQYLIT RAEAQLQVWI PPLNVRGGRG AIILLTCVVH PELIFDITKI
HCV_K1_S2 LIWWLQYFIT RAEACLQVWI PPLNVRGGRD AIILLTCALH PELIFSITKI
HCV_K1_S3 LIWWLQYFIT RAEAHLQVWI PPLNVRGGRD AIILLTCAVH SELVFEITKI
HCV_L2    LIWWLQYFIT RAEAHLQVWV PPLNVRGGRD AVILLTCAVY PELVFDITKL
HCV_N     LIWWLQYLIA RAEAHLHVWV PPLNVRGGRD AIILLTCAVH PELIFDITKL
HCV12083  FIWWNQYFLA RAEAMVHEWV PDLRVRGGRD SIILLTCLLH PQLGFEVTKI
HCV1480   VMWWLQYFIA RVEAIIHVWV PPLQVKGGRD AVIMLTCLFH PALGFEITKI
HCVPOLYP  LIWWLQYYIT RAEALLQVWI PPLNVRGGRD AIILLTCVVH PELIFDITKL
HD_1      LIWWLQYFIT RAEALLQVWI PPLNVRGGRD AIILLTCAVH PGLIFDITRL
HPCCGAA   CMWWLQYFLT RVEAQLHVWV PPLNVRGGRD AVILLTCVVH PALVFDITKL
HPCFG     LIWWNQYAIC RGEAALQVWV PPLLVRGSRD SVILLASLLY PSLIFDITKL
HPCGENANTI LIWWLQYFIT RAEAHLQVWI PPLNVRGGRD AIILLTCAAY PELIFDITKI
HPCGENOM  LIWWLQYFLT IAEAHLQVWI PPLNIRGGRD AIILLTCAIH PESIFDITKL
```

```
HPCHUMR   LIWWLQYFTT RAEADLHVWI PPLNARGGRD AIILLMCAVH PELIFDITKL
  HPCJ    LIWWLQYLLT RAEAHLHMWI PPLNVRGGRD AIILLACAAH PELIFDITKL
 HPCJCG   LIWWLQYFIT RAEAHLQVWV PPLNVRGGRD AIILLTCAVH PELIFDITKL
HPCJK046  FLWWTQYFIA HIEAKLHVWV PFLRVRGGRD AIILLTCVFH PSLGFEVTKI
HPCJK049  LIWWNQYTIA RCEAALQIWV PPLLARGARD GIILLAGLFY PALVFDITKL
 HPCJTA   LIWWLQYFIT RAEAHLCVWV PPLNVRGGRD AIILLTCAAH PELIFDITKL
 HPCJTB   LIWWLQYFIT RAEAHLCVWV PPLNVRGGRD AIILLTCAAH PELIFDITKL
 HPCK3A   LIWWNQYTIC RCESALHVWV PPLLARGGRD GVILLTSLLY PSLIFDITKL
HPCPLYPRE CLWWLQYFLT RVEAQLHVWI PPLNVRGGRD AVILLMCAVH PTLVFDITKL
 HPCPOLP  FLWWLCYLLT LAEAMVQEWA PPMQVRGGRD GIIWAVAIFC PGVVFDITKW
  HPCPP   LIWWLQYLIT RAEAHLQVWI PPLNVRGGRD AIILLTCAVH PELIFDITKL
HPCUNKCD  FIWWLQYLIT RTEAHLQVWV PPLNVRGGRD AIILLTCVVH PELIFDITKY
  MKC1A   LIWWLQYLIT RAEAHLQVWI PPLNVRGGRD AIILLTCAVH PELIFDITKL
  NDM59   SLWWLCYLLT LGEAMIQEWV PPMQARGGRD GIIWAATMFC PGVVFDITKW
   NZLI   LMWWNQYTIC RCESALHVWV PPLLARGSRD GVILLTSLLY PSLIFDITKL
   SA13   VIWWLQYFIA RVEAFIHVWV PPLQVRGGRD AIIMLTCLFH PALGFEVTKI
  Th580   AIWWNQYMLA RAEAMIHDWV PDLRVRGGRD AIILLTCLLH PHLGFEVTKI
Type_3a_CB LIWWNQYTIC RCESALQVWV PPLLARGSRD GVILLTSLLY PSLIFDITKL
 TypeV_D  LIWWNQYTIC RCESALQVWV PPLLARGSRD GVILLTSLLY PSLIFDITKL
  VN004   MLWWIQYFIA LTEAQLHQWV PSLLVRGGRD AVILLACLFH PQLGFEVTKI
  VN235   FLWWNQYFIA RSEALIQQWV PSLRVRGGRD AVILLTCLLH PSLGFDITKM
  VN405   TLWWIQYYIA RAEAMLYVWV PSLQVRGGRD AVILLTCLLH PQLGFEVTKA

          901                                                 950
  BEBE1   LLAILGPGYL FRAAVMRTPY FVRANALLRM CALVKQLAGG KYVQVALITL
  D89815  LLSILGPLMV LQASLIRVPY FVRAQGLIRA CMLVRKAAGG HYVQMAFVKL
ED43type_4 LLAILGPLHI LQASLLRIPY FVRAQALVKI CSLLRGVVYG KYFQMVVLKS
  HC_C2   LLAIFGPIMV LQAGITKMPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMKL
  HC_G9   MLAILGPLWI LQASLLKVPY FVRAHGLIRL CMLVRKTAGG QYVQMALLKL
 HCU16326 LLAIFGPLMV LQAGITRVPY FVRAQGLIRA CMLARKVVGG HYVQMVFMKL
 HCV_H_CMR LLAIFGPLWI LQASLLKVPY FVRVQGLLRI CALARKIAGG HYVQMAIIKL
  HCV_J1  LLAVLGPLWI LQASLLKVPY FVRVQGLLRI CALARKMVGG HYVQMAIIKL
 HCV_J483 LLAILGPLMV LQAGITRVPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMKL
  HCV_J8  LLAILGPAYL LKASLLRIPY FVRAHALLRV CTLVKHLAGA RYIQMLLITI
  HCV_JK1 LLAIFGPLMV LQAGLTRVPY FVRAQGLLRA CMLVRKVAGG HYVQMALMKL
  HCV_JS  LLAILGPLMV LQAGITRVPY FVRAQGLIRA CMLVRKVAGG HYVQMAFVKL
 HCV_K1_R1 VLAILGPLMV LQAGIAKVPD FVRAQGVIRE CILVRKVGGG QYVQMAFMKL
 HCV_K1_R2 LVAILGPLMV LQAGITKMPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMKL
 HCV_K1_R3 LLAIFGPLMV LQAGLTKVPY FVRAQGLIRA CMLVRKVAGG HYVQMAFVKL
 HCV_K1_S1 MLAILGPLMV LQAGIAKVPD FVRAQGVIRE CILVRKVGGG QYVQMAFMKL
 HCV_K1_S2 LVAILGPLMV LQAGITKMPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMKL
 HCV_K1_S3 LLAIFGPLMV LQAGLIRVPY FVRAQGLIRV CMLVRKVAGG HYVQMAFVKL
  HCV_L2  LLAIFGPLMV LQAGIIKMPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMKL
  HCV_N   LIAILGPLMV LQAGITRVPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMRL
 HCV12083 LLAVLAPLYI LQYSLLKVPY FVRAHILLRA CLLVRRLAGG KYVQACLLRL
 HCV1480  LFGILGPLYL LQHSLTKVPY FLRARALLRL CLLAKHLVYG KYVQAALLHL
 HCVPOLYP LLAMLGPPMV LQAVITKVPY FVRAQGLIRA CMLVRKVAGG HYVQMALMKL
   HD_1   LLAILGPLMV LQAGITRVPY FVRAHGLIRA CMLVRKVAGG HYVQMAFMKL
 HPCCGAA  LLAIFGPLWI LQASLLKVPY FVRVQGLLRI CALARKIAGG HYVQMAIIKL
  HPCFG   LIAVLGPLYL IQAALTSTPY FVRAHVLIRI CMLVRSAMGG KYVQMAVLTV
HPCGENANTI LLAILGPLMV LQAGLTRIPY FVRAQGLIRA CMLVRKAAGG HYVQMALMKL
 HPCGENOM LLATLGPLLV LQAGITRVPY FVRAHGLIRA CMLLRKVAGG HYVQMAFMKL
 HPCHUMR  LIAILGPLMV LQAGITRVPY FVRAQGLIHA CMLVRKVAGG HYVQMAFMKL
  HPCJ    LLAILGPLYM LQAGITAVPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMKL
 HPCJCG   LLAILGPLMV LQAGITRVPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMKL
HPCJK046  LLALIGPLYL LHASLLRVPY YVRAHALIRI CALVQNVAGG KYVQAAILRA
HPCJK049  LLAILGPLYI LQASLVRVPY FVRAHAVVRL CILVRNITGG KYVQMVLLAL
 HPCJTA   LLAILGPLMV LQAAITAMPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMKL
 HPCJTB   LLAILGPLMV LQAAITAMPY FVRAQGLIRA CMLVRKVAGG HYVQMAFMKL
 HPCK3A   LIAALGPLYL IQATITATPY FVRAHVLVRL CMLVRSVMGG KYFQMIILSL
HPCPLYPRE LLAVFGPLWI LQASLLKVPY FVRVQGLLRF CALARKMIGG HYVQMVIIKL
```

```
HPCPOLP   LLAVLGPAYL LKGALTRVPY FVRAHALLRM CTMVRHLAGG RYVQMVLLAL
  HPCPP   LLAILGPLMV LQAGLTQMPY FVRAQGLIRM CMLVRKAAGG HYVQMALMKL
HPCUNKCD  LLAIFGPLMV LQAGITRVPY FVRAQGLIRA CMLARKVVGG HYVQMVFMKL
  MKC1A   LLAILGPLMV LQAGITRVPY FVRAQGLIRA CMLVWKAAGG HYVQMAFVKL
  NDM59   LLAMLGPAYL LRDALTRVPY FVRAHALLRL CTMVRHLAGG RYVQMALLAL
    NZLI  LMAVLGPLYL IQATITTTPY FVRAHVLVRL CMLVRSVIGG KYFQMIILSI
    SA13  LLGILGPLYL LQYSLIKLPY FIRARALLRA CLLAKHLACG RYVQAALLHL
   Th580  LLAILAPLYI LQHSLLKVPY FVRAHILLRA CMFFRKVAAG KYVQACLLRL
Type_3a_CB LIAVLGPLYL IQAAVTTTPY FVRAQVLVRL CMLVRSVMGG KYFQMIILSI
  TypeV_D LIAVLGPLYL IQASITATPY FVRAHVLVRL CMLVRSMMGG KYFQMIILSV
   VN004  LLALLGPLYL LQYSLLKTPY FVRAHILLRA CMFFRGMARG RYAQAILLRI
   VN235  LLALLGPLYL LQVSLLRVPY YVRAHALLRV CILVRRVAGG KYIQAALLKL
   VN405  ILALLGPLYI LQYSLLKTPY FVRAHILLRV CMFLRGVAGG KYVQAALLRL

          951                                                1000
   BEBE1  GKWTGTYIYD HLSPMSDWAA DGLRDLAVAV EPIVFSPMER KVIVWGAETT
   D89815 AALTGTYVYD HLTPLQDWAH VGLRDLAVAV EPVVFSAMET KVITWGADTA
ED43type_4 RGLTGTYIYD HLTPMSDWPP YGLRDLAVAL EPVVFTPMEK KVIVWGADTA
   HC_C2  AALTGTYVYD HLTPLRDWAH TGLRDLAVAV EPVVFSDMET KIITWGADTA
   HC_G9  GAFAGTYIYN HLSPLQDWAH SGLRDLAVAT EPVIFSRMEI KTITWGADTA
  HCU16326 AALAGTYVYD HLTPLRDWAH TGLRDLAVAV EPVVFSDMET KVITWGADTA
 HCV_H_CMR GALTGTYVYN HLTPLRDWAH NGLRDLAVAV EPVVFSRMET KLITWGADTA
   HCV_J1  GALTGTYVYN HLTPLRDWAH NGLRDLAVAV EPVVFSQMET KLITWGADTA
  HCV_J483 GALTGTYVYN HLTPLRDWAH AGLRDLAVAV EPVVFSDMET KVITWGADTA
   HCV_J8  GRWTGTYIYD HLSPLSTWAA QGLRDLAIAV EPVVFSPMEK KVIVWGAETV
  HCV_JK1  AALTGTYVYD HLTPLRDWAH SGLRDLVVAV EPVVFSDMET KIITWGADTA
  HCV_JS   AALTGTYVYD HLAPLQHWAH TGLRDLAVAV EPVVFSDMET KIITWGADTA
 HCV_K1_R1 AELKGTSVYD HLTPLRDWAH TGLRDLAVAV EPVVFSDMET KIITWGADTA
 HCV_K1_R2 AALTGTYVYD HLTPLQDWAH TGLRDLAVAV EPVVFSDMEI KIITWGGDTA
 HCV_K1_R3 AALTGTYVYN HLTPLQDWAH TGLRDLAVAV EPVVFSDMET KIITWGADTA
 HCV_K1_S1 AELKGTSVYD HLTPLQDWAH TGLRDLAVAV EPVVFSDMET KIITWGADTA
 HCV_K1_S2 AALTGTYVYD HLTPLQDWAH TGLRDLAVAV EPVVFSDMEI KIITWGGDTA
 HCV_K1_S3 AALTGTYVYN HLTPLQDWAH AGLRDLAVAV EPVVFSDMET KIITWGADTA
   HCV_L2  AALTGTYVYD HLTPLRDWAH TGLRDLAVAV EPVVFSDMET KIITWGADTA
   HCV_N   GALTGTYVYN HLTPLRDWAH AGLRDLAVAV EPVTFSDMET KIITWGADTA
  HCV12083 GAWTGTFVYD HLAPLSDWAS DGLRDLAVAV EPVIFSPMEK KIITWGADTA
  HCV1480  GRLTGTYIYD HLAPMKDWAA SGLRELTVAT EPIVFSAMET KVITWGADTA
  HCVPOLYP AGLTSTYVYD HLTPLQDWAH GGLRDLAVAV EPVVFSDMET KIITWGADTA
    HD_1   AALTGTYVYD HLTPLRDCRH AGLRDLAVAV EPVVFSAMET KIITWGADTA
  HPCCGAA  GALTGTCVYN HLAPLRDWAH NGLRDLAVAV EPVVFSRMET KLITWGADTA
   HPCFG   GRWFNTYLYD HLSPIQDWAA EGLKGLAVAT EPVIFSPMEI KVITWGADTA
HPCGENANTI AALTGTYVYD HLTPLQDWAH TGLRDLAVAV EPVVFSDMET KIITWGADTA
 HPCGENOM  GALTGTYVYN HLTPLQYWPR AGLRELAVAV EPVIFSDMET KIITWGADTA
 HPCHUMR   GALTGTYIYN HLTPLRDWPR AGLRDLAVAV EPVVFSDMET KIITWGADTA
   HPCJ    AALTGTYVYD HLTPLRDWAH AGLRDLAVAV EPVVFSDMET KIITWGADTA
  HPCJCG   AALTGTYVYD HLTPLRDWAH AGLRDLAVAV EPVVFSDMET KLITWGADTA
 HPCJK046  GSWTGTYIYD HLVPLRTWAS DGLRDLAVAV EPVVFSPMEK KVITWGADTA
 HPCJK049  ARGFNTYLYD HLSPMTDWAA EGLKDLAVAV EPVIFSPMEV KVITWGADTT
  HPCJTA   AALTGTYVYD HLTPLQDWAH AGLRDLAVAV EPVVFSDMET KIITWGADTA
  HPCJTB   AALTGTYVYD HLTPLQDWAH AGLRDLAVAV EPVVFSDMET KIITWGADTA
  HPCK3A   ADGSNTYLYD HLAPMQHWAA AGLKDLAVAT EPVIFSPMEI KVITWGADTA
HPCPLYPRE  GALTGTYVYN HLTPLRDWAH NGLRDLAVAV EPVVFSQMET KLITWGADTA
 HPCPOLP   GRWTGTYIYD HLTPMSDWAA NGLRDLAVAV EPIIFSPMEK KVIVWGAETA
   HPCPP   AALTGTYVYD HLTPLQDWAH AGLRDLAVAV EPVVFSDMET KIITWGAETA
 HPCUNKCD  AALAGTYVYD HLTPLRDWAH TGLRDLAVAV EPVVFSDMET KVITWGADTA
  MKC1A    AALTGTYIYD HLTPLRDWAH AGLRDLAVAV EPVVFSDMET KIITWGADTA
  NDM59    GRWTGTYIYD HLSPMSDWAA SGLRDLAVAV EPIIFSPMEK KVIVWGAETA
    NZLI   GRWFNTYLYD HLAPMQHWAA AGLKDLAVAT EPVIFSPMEI KVITWGADTA
    SA13   GRLTGTYIYD HLAPMKDWAA SGLRDLAVAT EPIIFSPMET KVITWGADTA
   Th580   GAWTGTYIYD HLAPLSEWAS DGLRDLAVAV EPVIFSPMEK KIITWGADTA
Type_3a_CB GRWFNTYLYD HLAPMQHWAA AGLKDLAVAT EPVIFSPMEI KVITWGADTA
```

```
TypeV_D   GRWFNTYLYD HLAPMQHWAA AGLKDLAVAT EPVIFSPMEI KVITWGADTA
VN004     GAWTGTYIYD HLAPLSDWAC DGLRDLAVAV EPVVFSPMEK KVITWGADTA
VN235     GAWTGTYIYD HLAPLSTWAS DGLRDLAVAV EPVTFSPMEK KIITWGADTA
VN405     GAWTGTYIYD HLTPLSDWAC DGLRDLAVAV EPVVFSPMEK KVITWGADTV

          1001                                                1050
BEBE1     ACGDIIHGLP VSARLGQEVL LGPADGYTSK GWRLLAPITA YAQQTRGLLS
D89815    ACGDIISGLP VSARRGKEIL LGPADSFEGQ GWRLLAPITA YSQQTRGLLG
ED43type_4 ACGDIIRGLP VSARLGNEIL LGPADTETSK GWRLLAPITA YAQQTRGLFS
HC_C2     ACGDIILGLP VSARRGKEIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
HC_G9     ACGDIINGLP VSARRGREVL LGPADALTDK GWRLLAPITA YAQQTRGLLG
HCU16326  ACGDIILACP ASARRGKEIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
HCV_H_CMR ACGDIINGLP VSARRGQEIL LGPADGMVSK GWRLLAPITA YSQQTRGLLG
HCV_J1    ACGDIINGLP VSARKGREIL LGPADGMVSK GWRLLAPITA YAQQTRGLLG
HCV_J483  ACGDIILGLP VSARRGKEIF LGPADSLEGQ GWRLLAPITA YAQQTRGLLG
HCV_J8    ACGDILHGLP VSARLGREVL LGPADGYTSK GWKLLAPITA YSQQTRGVLG
HCV_JK1   ACGDIILGLP VSARRGREIL LGPADGFREQ GWRLLAPITA YTQQTRGLLG
HCV_JS    ACGDIILGLP VSARRGREIL LGPADSLEGQ GWRLLAPITA YSQQTRGLFG
HCV_K1_R1 ACGDIIMGLP VSARRGREIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
HCV_K1_R2 ACGDIILGLA VSARRGKEIF LGPADSLEGS GWRLLAPITA YSQQTRGLFG
HCV_K1_R3 ACGDIISGLA VSARRGREIL LGPADSFEGQ GWRLLAPITA YSQQTRGLLG
HCV_K1_S1 ACGDIIMGLP VSARRGREIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
HCV_K1_S2 ACGDIILGLA VSARRGKEIF LGPADSLEGS GWRLLAPITA YSQQTRGLFG
HCV_K1_S3 ACGDIISGLP VSARRGREIL LGPADSFEGQ GWRLLAPITA YSQQTRGLLG
HCV_L2    ECGDIILGYR SSARRGREIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
HCV_N     ACGDIILGLP VSARRGREIL LGPADSLVGQ GWRLLAPITA YAQQTRGLLG
HCV12083  ACGDILSGLP VSARLGNLVL LGPADDMQRG GWKLLAPITA YSRQTRGLLG
HCV1480   ACGNILAVLP VSARRGREIF LGPADDIKTS GWRLLAPITA YAQQTRGLVG
HCVPOLYP  ACGDIISGLP VSARRGREIL LGPADSLKEQ GWRLLAPITA YAQQTRGVLG
HD_1      ACGDIILGLP VSARRGREIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
HPCCGAA   ACGDIINGLP VSARRGQEIL LGPADGMVSK GWRLLAPITA YAQQTRGLLG
HPCFG     ACGDILCGLP VSARLGRELL LGPADDYKKM GWRLLSPISA YAQQTRGLLG
HPCGENANTI ACGDIILGLP VSARRGREIL LGPADSLEGR GWRLLAPITA YAQQTRGLFG
HPCGENOM  ACGDIILGLP VSARRGKEIL LGPADSLEGQ GWRLLAPITA YAQQTRGLFG
HPCHUMR   ACGDIILGLP VSARRGKEIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
HPCJ      ACGDIILGLP VSARRGREIL LGPADSLEGR GLRLLAPITA YSQQTRGLLG
HPCJCG    ACGDIISGLP VSARRGKEIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
HPCJK046  ACGDILAGLP VSARRGNLIL LGPADDVKDK GWSLLAPITA YSQQTRGLLG
HPCJK049  ACGDILCGLP VSARLGKEVL LGPADDYRSM GWRLLAPITA YAQQTRGLLG
HPCJTA    ACGDIILGLP VSARRGREIL LGPADSIEGQ GWRLLAPITA HAQQTRGLFG
HPCJTB    ACGDIILGLP VSARREREIL LGPADSIEGQ GWRLLAPITA YAQQTRGLLG
HPCK3A    ACGDILCGLP VSARLGREVL LGPADDYREM GWRLLAPITA YXQQTRGLLG
HPCPLYPRE ACGDIINGLP VSARRGREIL LGPADGMVSK GWRLLAPITA YAQQTRGLLG
HPCPOLP   ACGDILHGLP VSARLGREVL LGPADGYTSK GWSLLAPITA YAQQTRGLLG
HPCPP     ACGDIISSLP VSARRGRELL LGPADSFDGQ GWRLLAPITA YSQQTRGLLG
HPCUNKCD  ACGDIILALP ASARRGKEIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
MKC1A     ACGDIILGLP VSARRGKEIL LGPADSLEGQ GWRLLAPITA YSQQTRGLLG
NDM59     ACGDVLHGLP VSARLGREIL LGPADGYTSK GWRLLAPITA YSQQTRGLLG
NZLI      ACGDILCGLP VSARLGREVL LGPADDYREM GWRLLAPITA YAQQTRGLLG
SA13      ACGDILAGLP VSARRGHEIF LGPADDIREA GWRLLAPITA YAQQTRGLLG
Th580     ACGDILRGLP VSARLGDLVL LGPADDMRHG GWKLLAPITA YAQQTRGVLG
Type_3a_CB ACGDILCGLP VSARLGREVL LGPADDYREM GWRLLAPITA YAQQTRGLVG
TypeV_D   ACGDILCGLP VSARLGREVL LGPADDYREM GWRLLAPITA YAQQTRGLLG
VN004     ACGDIIAGLP VAARRGNLLF LGPADDVKGK GWRLLAPITA YAQQTRGLLG
VN235     ACGDILAGLP VSARLGHLLF LGPADDMKSM GWRLLAPITA YAQQTRGIVG
VN405     ACGDIISGLP VSARRGNLIF LGPADDIRDG GWRLLAPITA YCQQTRGLLG

          1051                                                1100
BEBE1     AIVVSMTGRD KTDQAGEIQV LSTVTQSFLG TSISGVLWTV YHGAGNKTLA
D89815    CIITSLTGRD KNQVEGEVQV VSTAKQSFLA TCVNGACWTV FHGAGSKTLA
ED43type_4 TIVTSLTGRD TNENCGEVQV LSTATQSFLG TAVNGVMWTV YHGAGAKTIS
```

```
          HC_C2  CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
          HC_G9  CIITSLTGRD KNQVEGEVQI VSTATQTFLA TCVNGVCWTV YHGAGSRTIA
        HCU16326 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCINGVCWTV FHGAGSKTLA
       HCV_H_CMR CIITSLTGRD KNQVEGEVQI VSTATQTFLA TCINGVCWTV YHGAGTRTIA
          HCV_J1 CIITSLTGRD KNQVEGEVQI VSTAAQTFLA TCINGVCWTV YHGAGTRTIA
        HCV_J483 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCINGVCWTV YHGAGSKTLA
          HCV_J8 AIVVSLTGRD KNEQAGQVQV LSSVTQTFLG TSISGVLWTV YHGAGNKTLA
         HCV_JK1 CIVTSLTGRD KNQVEGEAQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
          HCV_JS CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCINGVCWTV FHGAGSKTLA
        HCV_K1_R1 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
        HCV_K1_R2 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCTNGVCWTV YHGAGSKTLA
        HCV_K1_R3 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
        HCV_K1_S1 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
        HCV_K1_S2 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCTNGVCWTV YHGAGSKTLA
        HCV_K1_S3 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
          HCV_L2 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV FHGAGSKTLA
           HCV_N CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
        HCV12083 TIVTSLTGRD KNEVEGEVQV VSTDTQSFVA TSINGVMWTV YHGPGFKTLA
         HCV1480 AIVLSLTGRD KNEAEGEVQF LSTATQTFLG ICINGVMWTL FHGAGSKTLA
        HCVPOLYP CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
            HD_1 CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
         HPCCGAA CIITSLTGRD KNQVEGEVQI VSTATQTFLA TCINGVCWTV YHGAGTRTIA
           HPCFG TIVTSLTGRD KNVVTGEVQV LSTATQTFLG TTVGGVMWTV YHGAGSRTLA
      HPCGENANTI CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCINGVCWTV YHGAGSKTLA
        HPCGENOM CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCINGVCWTV YHGAGSKTLA
        HPCHUMR CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV YHGAGSKTLA
           HPCJ CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCINGVCWTV YHGAGSKTLA
         HPCJCG CIITSLTGRD KNQVDGEVQV LSTATQSFLA TCVNGVCWTV YHGAGSKTLA
       HPCJK046 TIVTSLTGRD KNEAAGEIQI LSTATQTFLA TCVNGVCWTV YHGAGSKTLA
       HPCJK049 TIVTSLTGRD KNIVTGEIQV LSTSTQTFLG TSVGGVMWTV YHGAGSRTLA
         HPCJTA CIVTSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV FHGAGSKTLA
         HPCJTB CIVTSLTGRD KNQVEGEVQV VSTATQSFLA TCVNGVCWTV FHGAGSKTLA
         HPCK3A TIVTSLTGRD KNVVAGEVQV LSTATQTFLG TTVGGVMWTV YHGAGSRTLA
      HPCPLYPRE CIITSLTGRD KNQVEGEVQI VSTAAQTFLA TCINGVCWTV YHGAGTRTIA
        HPCPOLP TIVVSMTGRD KTEQAGEIQV LSTVTQSFLG TTISGVLWTV YHGAGNKTLA
           HPCPP CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCINGVCWTV FHGAGSKTLA
       HPCUNKCD CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCINGVCWTV FHGAGSKTLA
          MKC1A CIITSLTGRD KNQVEGEVQV VSTATQSFLA TCINGVCWTV YHGAGSKTLA
          NDM59 AIVVSMTGRD KTEQAGEIQV LSTVTQSFLG TSISGVLWTV YHGAGNKTLA
           NZLI TIVTSLTGRD KNVVTGEVQV LSTATQTFLG TTVGGVIWTV YHGAGSRTLA
           SA13 AIIVSLTGRD KNEAEGEVQV LSTATQTFLG TCINGVMWTV FHGAGAKTLA
           Th580 TIVTSLTGRD KNEAEGEVQV VSTATQSFLA TTINGVLWTV YHGAGSKNLA
      Type_3a_CB TIVTSLTGRD KNVVTGEVQV LSTATQTFLG TTVGGVTWTV YHGAGSRTLA
         TypeV_D TIVTGLTGRD KNVVTGEVQV LSTATQTFLG TTVGGVMWTV YHGAGSRTLA
           VN004 TIVTSLTGRD KNEVEGEIQV VSTATQSFLA TAVNGVLWTV YYGAGSKTLA
           VN235 TIVTSLTGRD RNVVEGEIQV LSTATQSFLG TAINGVMWTV YHGAGSKTLA
           VN405 TIVTSLTGRD KNEVEGEIQV VSTATQSFLA TTVNGVLWTV YHGAGSKTLA

                 1101                                                1150
          BEBE1  GSRGPVTQMY SSAEGDLVGW PSPPGTRSLE PCTCGAVDLY LVTRNADVIP
          D89815 AAKGPITQMY TNVDQDLVGW PAPPGARSLT PCTCGSSDLY LVTRHADVIP
      ED43type_4 GPKGPVNQMY TNVDQDLVGW PAPPGVRSLA PCTCGSADLY LVTRHADVIP
          HC_C2  GPKGPITQMY TNVDQDLVGW PAPSGARSLT PCTCGSSDLY LVTRYADVIP
          HC_G9  SASGPVIQMY TNVDQDLVGW PAPQGARSLT PCTCGASDLY LVTRHADVIP
        HCU16326 GPKGPITQMY TNVDQDLVGW PAPPGARSLT PCTCGSSDLY LVTRHADVIP
       HCV_H_CMR SPKGPVIQMY TNVDQDLVGW PAPQGSRSLT PCTCGSSDLY LVTRHADVIP
          HCV_J1 SPKGPVIQMY TNVDQDLVGW PAPQGARSLT PCTCGSSDLY LVTRHADVIP
        HCV_J483 GPKGPITQMY TNVDLDLVGW QAPPGARSMT PCTCGSSDLY LVTRHADVIP
          HCV_J8 GPKGPVTQMY TSAEGDLVGW PSPPGTKSLD PCTCGAVDLY LVTRNADVIP
         HCV_JK1 GPKGPINQMY TNVDQDLVGW QAPSGAASLT PCTYGSSDLY LVTRHADVIP
          HCV_JS GPKGPITQMY TNVDQDLVGW QAPPGARSMT PCTCGSSDLY LVTRHADVIP
```

```
HCV_K1_R1   GPKGPITQMY TNVDQDLVGW PAPPGARSLT PCTCGGSDLY LVTRHADVIP
HCV_K1_R2   GPKGPITQMY TNVDQDLVGW QAPSGARSLT PCTCGSSDLY LVTRHADVIP
HCV_K1_R3   GPKGPITQMY TNVDQDLVGW QAPPGARSLT PCTCGSSDLY LVTRHADVIP
HCV_K1_S1   GPKGPITQMY TNVDQDLVGW PAPPGARSLT PCTCGGSDLY LVTRHADVIP
HCV_K1_S2   GSKGPITQMY TNVDQDLVGW QAPSGARSLT PCTCGSSDLY LVTRHADVIP
HCV_K1_S3   GPKGPITQMY TNVDQDLVGW QAPPGARSLT PCTCGSSDLY LVTRHADVIP
   HCV_L2   GPKGPITQMY TNVDQDLVGW QAAPGMRSLT PCTCGSSDLY LVTRHADVIP
    HCV_N   GPKGPIAQMY TNVDQDLVGW PAPPGARSLT PCTCGSSDLY LVTRHADVIP
 HCV12083   GPKGPVCQMY TNVDLDLVGW PSPPGARSLT PCNCGSSDLY LVTREADVIP
  HCV1480   GPKGPVVQMY TNVDKDLVGW PSPPGKGSLT RCTCGSADLY LVTRHADVIP
 HCVPOLYP   GPKGPVTQMY TNVDQDLVGW PAPPGARSLT PCTCGSSDLY LVTRHADVIP
     HD_1   GPKGPVTQMY TNVDQDLVGW QAPPGARSLT PCTCGSSDLY LVTRHADVIP
  HPCCGAA   SPKGPVIQTY TNVDQDLVGW PAPQGSRSLT PCTCGSSDLY LVTRHADVIP
    HPCFG   GNKRPALQMY TNVDQDLVGW PAPAGTKSLD PCTCGSSDLY LVTREADVLP
HPCGENANTI  GPKGPITQMY TNVDQDLVGW HAPQGARSLT PCTCGSSDLY LVTRHADVIP
  HPCGENOM  GPKGPITQMY TNVDQDLVGW HRPPGARSLT PCTCGSSDLY LVTRHADVIP
   HPCHUMR  APKGPITQMY TNVDQDLVGW PKPPGARSLT PCTCGSSDLY LVTRHADVIP
     HPCJ   GPKGPITQMY TNVDQDLVGW QAPPGARSMT PCTCGSSDLY LVTRHADVIP
    HPCJCG  GPKGPITQMY TNVDQDLVGW PAPPGARSMT PCTCGSSDLY LVTRHADVVP
  HPCJK046  GPRGPVCQMY TNVDQDMVGW PAPAGTRSYT PCTCGASDLY LITRQADVIP
  HPCJK049  GNKRPALQMY TNVDQDLVGW PSPPGAKSLV PCTCGSADLY LITRDADVLP
    HPCJTA  GPKGPITQMY TNVDQDLVGW HAPPGARSLT PCTCGSSDLY LVTRHADVIP
    HPCJTB  GPKGPITQMY TNVDQDLVGW HAPPGARSLT PCTCGSSDLY LVTRHADVIP
    HPCK3A  GVKHPALQMY TNVDQDLVGW PAPPGAKSLE PCTCGSADLY LVTRDADVIP
 HPCPLYPRE  SPKGPVIQMY TNVDQDLVGW PAPQGSRSLT PCTCGSSDLY LVTRHADVIP
   HPCPOLP  GSRGPVTQMY SSAEGDLVGW PSPPGTKSLE PCTCGAVDLY LVTRNADVIP
     HPCPP  GPKGPITQMY TNVDQDLVGW SAPPGARSLT PCTCGSSDLY LVTRHADVIP
  HPCUNKCD  GPKGPITQMY TNVDQDLVGW PAPPGARSLT PCTCGSSDLY LVTRHADVIP
     MKC1A  GPKGPITQMY TNVDQDLVGW PAPPGARSMT PCTCGSSDLY LVTRHADVIP
     NDM59  GSRGPVTQMY SSAEGDLVGW PSPPGTKSLE PCTCGAVDLY LVTRNADVIP
      NZLI  GAKHPALQMY TNVDQDLVGW PAPPGAKSLE PCACGSSDLY LVTRDADVIP
      SA13  GPKGPVVQMY TNVDKDLVGW PTPPGTRSLT PCTCGSADLY LVTRHADVVP
     Th580  GPKGPVCQMY TNVDQDLVGW PAPLGARSLA PCTCGSSDLY LVTRGADVIP
 Type_3a_CB GAKHPALQMY TNVDQDLVGW PAPPGAKSLE PCACGSADLY LVTRDADVIP
   TypeV_D  GAKHPALQMY TNVDQDLVGW PAPPGAKSLE PCACGSADLY LVTRDADVIP
     VN004  GPKGPVCQMY TNVDQDLVGW PAPAGARSLT PCSCGSSDLY LVTRNADVIP
     VN235  GPKGPVCQMY TNVDQDMVGW PAPPGTRSLT PCTCGASDLY LVTRNADVIP
     VN405  GPKGPICQMY TNVDQDLVGW PAPPGARSLT PCTCGSSDLY LVTRNADVIP

            1151                                              1200
    BEBE1   ARRRGDRRGA LLSPRPLSSL KGSSGGPVLC PRGHAVGIFR AAVCSRGVAK
   D89815   VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PSGHVVGIFR AAVCTRGVAK
ED43type_4  VRRRGDTRGA LLSPRPISIL KGSSGGPLLC PMGHRAGIFR AAVCTRGVAK
    HC_C2   VRRRGDSRGS LLSPRPVSYL KGSSGGPLLC PSGHAAGIFR AAVCTRGVAK
    HC_G9   VRRRGDNRGS LLSPRPISYL KGSSGGPLLC PMGHAVGIFR AAVCTRGVAK
 HCU16326   VRRRGDGRGS LLPPRPVSYL KGSSGGPLLC PSGHAVGILP AAVCTRGVAM
 HCV_H_CMR  VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PAGHAVGLFR AAVCTRGVTK
    HCV_J1  VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PAGHVVGIFR AAVCTRGVAK
  HCV_J483  VRRRGDSRGS LLSPRPVSYL KGSSGGPLLC PSGHVVGVFR AAVCTRGVAK
    HCV_J8  VRRKDDRRGA LLSPRPLSTL KGSSGGPVLC SRGHAVGLFR AAVCARGVAK
   HCV_JK1  VRRRGDSRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
    HCV_JS  VRRRGDSRGS LLSPRPVSYL KGSSGGPLPC PSGHAVGIFR AAVCTRGVAK
 HCV_K1_R1  VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PSGHVVGIFR AAVCTRGVAK
 HCV_K1_R2  VRRRGDGRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
 HCV_K1_R3  VRRRGDSRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
 HCV_K1_S1  VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PSGHVVGIFR AAVCTRGVAK
 HCV_K1_S2  VRRRGDGRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
 HCV_K1_S3  VRRRGDSRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
    HCV_L2  VRRRGDGRGS LLSPRPVSYL KGSSGGPLLW PSGHAVGIFR AAVCTRGVAK
     HCV_N  VRRRGDNRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGVFR AAVCTRGVAK
  HCV12083  ARRRGDSRAA LLSPRPISTL KGSSGGPIMC PSGHVVGLFR AAVCTRGVAK
```

```
HCV1480     ARRRGDTRAS LLSPRPISYL KGSSGGPIMC PSGHVVGVFR AAVCTRGVAK
HCVPOLYP    VRRRGDSRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
HD_1        VRRRGDSRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
HPCCGAA     VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PTGHAVGLFR AAVCTRGVAK
HPCFG       ARRRGDSTAS LLSTRPLSCL KGSSGGPVMC PSGHVVGIFR AAVCTRGVAK
HPCGENANTI  VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PSGHVVGIFR AAVCTRGVAK
HPCGENOM    VRRRGDSRGS LLSPRPVSYL KGSSGGPLLC PFGHVAGIFR AAVCTRGVAK
HPCHUMR     VRRRGDSRGS LLSPRPVSYL KGSSGGPLLC PFGHAVGIFR AAVCTRGVAK
HPCJ        VRRRGDGRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
HPCJCG      VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PSGHVVGIFR AAVCTRGVAK
HPCJK046    ARRRGDNRAG LISPRPISTL KGSSGGPLLC PSGHVVGLFR AAVCTRGVAK
HPCJK049    ARRRGDSTAS LLSPRPLACL KGSSGGPIMC PSGHVAGIFR AAVCTRGVAK
HPCJTA      VRRRGDGRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
HPCJTB      VRRRGDGRGS LLSPRPVSYL KGSSGGPLLC PSGHAVGIFR AAVCTRGVAK
HPCK3A      ARRRGDSTAS LLSPRPLARL KGSSGGPVMC PSGHVAGIFR AAVCTRGVAK
HPCPLYPRE   VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PAGHAVGIFR AAVCTRGVAK
HPCPOLP     ARRRGDKRGA LLSPRPLSTL KGSSGGPVLC PRGHAVGVFR AAVCSRGVAK
HPCPP       VHRRGDSRGS LLSPGPISYL KGSSGGPLPC PSGRVVGIFR AAVCTRGVAK
HPCUNKCD    VRRRGDGRGS LLPPRPVSYL KGSSGGPLLC PSGHAVGILP AAVCTRGVAM
MKC1A       VRRRGDSRGS LLSPRPISYL KGSSGGPLLC PSGHVVGIFR AAVCTRGVAK
NDM59       ARRRGDKRGA LLSPRPLSTL KGSSGGPVLC PRGHAVGIFR AAVCSRGVAK
NZLI        ARRRGDSTAS LLSPRPLACL KGSSGGPVMC PSGHVAGIFR AAVCTRGVAK
SA13        ARRRGDTRAS LLSPRPISYL KGSSGGPVMC PSGHVVGVFR AAVCTRGVAK
Th580       ARRRGDTRAA LLSPRPISTL KGSSGGPLMC PSGHVVGLFR AAVCTRGVAK
Type_3a_CB  ARRRGDSTAS LLSPRPLACL KGSSGGPVMC PSGHVAGIFR AAVCTRGVAK
TypeV_D     ARRRGDSTAS LLSPRPLACL KGSSGGPVMC PSGHVAGIFR AAVCTRGVAK
VN004       ARRRGDNRAA LLSPRPISTL KGSSGGPMLC PSGHVAGIFR AAVCTRGVAK
VN235       ARRRGDTRAG LLSPRPLSTL KGSSGGPLMC PSDHVVGLFR AAVCTRGVAK
VN405       ARRRGDTRAA LLSPRPISTL KGSSGGPMLC PSGHVAGIFR AAVCTRGVAK

            1201                                                1250
BEBE1       SIDFIPVESL DVVTRSPNFT DNSTPPAVPQ TYQVGYLHAP TGSGKSTKVP
D89815      AVDFIPVESM ETTMRSPVFT DNSTPPAVPQ TFQVAHLHAP TGSGKSTKVP
ED43type_4  AVDFVPVESL ETTMRSPVFT DNSTPPAVPQ TYQVAHLHAP TGSGKSTKVP
HC_C2       AVDFVPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HC_G9       AVDFVPVESL ETTMRSPVFT DNSSPPTVPQ SYQVAHLHAP TGSGKSTKVP
HCU16326    AVEFIPVESM ETTMRSPVFT DNPSPPAVPQ TFQVAHLHAP TGSGKSTRVP
HCV_H_CMR   AVDFIPVENL ETTMRSPVFT DNSSPPAVPQ SFQVAHLHAP TGSGKSTKVP
HCV_J1      AVDFIPVESL ETTMRSPVFT DNSSPPAVPQ SFQVAHLHAP TGSGKSTKVP
HCV_J483    AVDFIPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HCV_J8      SIDFIPVESL DVATRTPSFS DNSTPPAVPQ SYQVGYLHAP TGSGKSTKVP
HCV_JK1     AVDFIPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HCV_JS      AVDFIPVESM ETTMRAPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HCV_K1_R1   AVDFVPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HCV_K1_R2   AVDFIPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSNKVP
HCV_K1_R3   AVDFVPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HCV_K1_S1   AVDFVPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HCV_K1_S2   AVDFIPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSNKVP
HCV_K1_S3   AVDFVPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HCV_L2      AVDFVPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HCV_N       AVDFVPVESM ETTMRSPVFT DTSSPPAVPQ TFQVAHLHAP TGSGKSTRVP
HCV12083    SLDFIPVENM ETTMRSPSFT DNSTPPAVPQ TYQVGYLHAP TGSGKSTRVP
HCV1480     ALEFVPVENL ETTMRSPVFT DNSTPPAVPH EFQVGHLHAP TGSGKSTKVP
HCVPOLYP    AVDFVPVESM ETTMRSPVFT DNSSPPAVPE TFQVAHLHAP TGSGKSTKVP
HD_1        AVDFVPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKTTKVP
HPCCGAA     AVDFIPVENL ETTMRSPVFT DNSSPPAVPQ SFQVAHLHAP TGSGKSTKVP
HPCFG       ALQFIPVETL STQVRSPSFS DNSTPPAVPE SYQVGYLHAP TGSGKSTKVP
HPCGENANTI  AVDFVPVESM ETTMRSPVFT DNSSPPAVPQ AFQVAHLHAP TGSGKSTKVP
HPCGENOM    AVDFIPVETM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HPCHUMR     AVDFVPVESM ETTMRSPVFT DNSSPPAVPQ SFQVAHLHAP TGSGKSTKVP
HPCJ        AVDFIPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
```

```
HPCJCG   AVDFIPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HPCJK046 ALDFVPCEAM DATTRSPTFT DNSTPPAVPQ AYQVGYLHAP TGSGKSTKVP
HPCJK049 ALQFIPVESL SAQTRSPSFS DNSTPPAVPQ TFQVGYLHAP TGSGKSTKVP
HPCJTA   AVDFIPVESM ETTMRSPVFT DNSSPPAVPQ TFQVGYLHAP TGSGESTKVP
HPCJTB   AVDFIPVESM ETTMRSPVFT DNSSPPAVPQ TFQVAHLHAP TGSGKSTKVP
HPCK3A   ALQFIPVETL STQARSPSFS DNSTPPAVPQ TFQVAHLHAP TGSGKSTKVP
HPCPLYPRE AVDFIPVENL ETTMRSPVFT DNSSPPVVPQ SYQVGYLHAP TGSGKSTKVP
HPCPOLP  SIDFIPVETL DIVTRSPTFS DNSTPPAVPQ SFQVAHLHAP TGSGKSTKVP
HPCPP    AVDFVPVESM ETTMRSPVFT DNSTPPAVPQ TYQVGYLHAP TGSGKSTKVP
HPCUNKCD AVEFIPVESM ETTMRSPVFT DNPSPPAVPQ TFQVAHLHAP TGSGKSTRVP
MKC1A    AVDFVPVESM ETTMRSPVFT DNSTPPAVPQ TFQVAHLHAP TGSGKSTRVP
NDM59    SIDFIPVEAL DVVTRSPTFS DNSTPPAVPQ SFQVAHLHAP TGSGKSTKVP
NZLI     SLQFIPVETL STQARSPSFS DNSTPPAVPQ TYQVGYLHAP TGSGKSTKVP
SA13     ALDFIPVENL ETTMRSPVFT DNSTPPAVPQ SYQVGYLHAP TGSGKSTKVP
Th580    ALDFIPVENM DTTMRSPVFT DNSTPPAVPH EFQVGHLHAP TGSGKSTKVP
Type_3a_CB ALQFIPVETL STQTRSPSFS DNSSPPAVPQ TYQVGYLHAP TGSGKSTRVP
TypeV_D  ALQFIPVETL STQARSPSFS DNSTPPAVPQ SYQVGYLHAP TGSGKSTKVP
VN004    SLDFAPVESM QSSQRSPSFS DNSTPPAVPQ TYQVGYLHAP TGSGKSTKVP
VN235    ALDFVPVENM ETTMRSPVFT DNSTPPAVPQ TYQVGYLHAP TGSGKSTKVP
VN405    SLDFVPVENM QSTARSPSFS DNTTPPAVPQ TYQVGYLHAP TGSGKSTKVP

         1251                                                1300
BEBE1    AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA YGINPNIRTG VRTVTTGDAI
D89815   AAYAAQGYMV LVLNPSVAAT LGFGAYMSKA HGIDPNIRTG VRTITTGAPI
ED43type_4 AAHAAQGYKV LVLNPSVAAT LGFGVYMSKA YGIDPNIRSG VRTITTGAPI
HC_C2    AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGIDPNTRTG VRTITTGAPI
HC_G9    AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGIDPNVRTG VRTITTGSPI
HCU16326 AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGIDPNLRTG VRTITTGAPI
HCV_H_CMR AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG VRTITTGSPI
HCV_J1   AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGIDPNIRTG VRTITTGSPI
HCV_J483 AAYAAQGYKV LVLNPSVAAT LGFGAYMPKA HGIDPNIRTG VRTITTGGPI
HCV_J8   AAYASQGYKV LVLNPSVAAT LGFGAYMSKA HGINPNIRTG VRTVTTGDSI
HCV_JK1  AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNISTG VRTITTGAPI
HCV_JS   AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA YGTDPNIRTG VRTITTGAPI
HCV_K1_R1 AAYAAQGYKV LVLNPSVTAT LGFGAYMSKA HGVDPNIRTG VRTITTGAPI
HCV_K1_R2 VEYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG VRTITTGAPI
HCV_K1_R3 AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG ARTITTGAPI
HCV_K1_S1 AAYAAQGYKV LVLNPSVTAT LGFGAYMSKA HGVDPNIRTG VRTITTGAPI
HCV_K1_S2 VEYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG VRTITTGAPI
HCV_K1_S3 AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG ARTITTGAPI
HCV_L2   AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGTDPNIRTG ARTITTGAPI
HCV_N    AAYATQGYKV LVLNPSVAAT LGFGAYMSKA HGTDPNIRTG VRTITTGAPI
HCV12083 AAYASQGYKV LVLNPSVAAT LSFGSYMRQA YGVEPNIRTG VRTVTTGGAI
HCV1480  AAYAAQGYKV LVLNPSVAAT FGFGAYMSRA YGVDPNIRTG VRTVTTGAGI
HCVPOLYP AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGTDPNIRTG VRTITTGAPI
HD_1     AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGIDPNIRTG VRTITTGAPI
HPCCGAA  AAYAAKGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG VRTITTGSPI
HPCFG    AAYVAQGYSV LVLNPSVAAT LGFGTYMSKA YGIDPNIRTG TRTITTGAKL
HPCGENANTI AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG VRTITTGAPI
HPCGENOM AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG VRTITTGAPI
HPCHUMR  AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGIDPNIRTG VRTITTGAPV
HPCJ     AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA YGTDPNVRTG VRTITTGSPI
HPCJCG   AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGIEPNIRTG VRTITTGGPI
HPCJK046 VAYASQGYKV LVLNPSVAAT LSFGSYLSRA HGIDPNIRTG VRTITTGAPI
HPCJK049 ASYVAQGYTV LVLNPSVAAT LGFGRFMSHA YGIDPNVRTG TRTITTGAKL
HPCJTA   AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGTDPNIRTG VRTITTGAPI
HPCJTB   AAYAAXGYKV LVLNPSVAAT LGFGAYMSKA HGTDPNIRTG VRTITTGGPI
HPCK3A   AAYVAQGYNV LVLNPSVAAT LGFGSFMSRA YGIDPNIRTG NRTVTTGAKL
HPCPLYPRE AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGIDPNIRTG VRTITTGSPI
HPCPOLP  VAYAAQGYKV LVLNPSVAAT LGFGAYLSKA HGINPNIRTG VRTVTTGAPI
HPCPP    AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG VRTITTGAPI
```

```
HPCUNKCD  AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGIDPNLRTG VRTITTGAPI
   MKC1A  AAYAAQGYKV LVLNPSVAAT LGFGAYMSKA HGVDPNIRTG VRTITTGAPI
   NDM59  VAYAAQGYKV LVLNPSVAAT LGFGAYLSKA HGINPNIRTG VRTVTTGESI
    NZLI  AAYVAQGYNV LVLNPSVAAT LGFGSFMSRA YGIDPNIRTG NRTVTTGAKL
    SA13  AAYAAQGYKV LVLNPSVAAT LGFGAYMSRA YGVDPNIRTG VRTVTTGAAI
    Th580 AAYATQGYKV LVLNPSVAAT LSFGAYMSKA HGIDPNIRTG VRTITTGGPV
Type_3a_CB AAYVAQGYTV LVLNPSVAAT LGFGSFMSRA YGIDPNIRTG NRTVTTGAKL
  TypeV_D  AAYVAQGYNV LVLNPSVAAT LGFGSFMSRA YGTDPNIRTG NRTVTTGAKL
    VN004  AAYAAQGYKV LVLNPSVAAT LGFGSYMSTS HGIDPNIRTG VRTITTGGAI
    VN235  AAYASQGYKV LVLNPSVAAT LGFGSYMSTA HGIDPNIRTG VRTITTGGPI
    VN405  AAYAAQGYKV LVLNPSVAAT LGFGSYMSTA HGIDPNIRTG VRTITTGGAI

          1301                                                1350
    BEBE1 TYSTYGKFLA DGGCSGGAYD VIICDECHSV DSTTILGIGT VLDQAETAGV
    D89815 TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTSILGIGT VLDQAETVGA
ED43type_4 TYSTYGKFLA DGGCSGGAYD IIICDECYST DSTTILGIGT VLDQAETAGV
    HC_C2 TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDQAETAGA
    HC_G9 THSTYGKFLA DGGCSGGAYD IIICDECHSV DATSILGIGT VLDQAETAGV
  HCU16326 TYSTYGKFLA DGGGSGGAYD IIMCDECHST DSTTIYGIGT VLDQAETAGA
 HCV_H_CMR TYSTYGKFLA DGGCSGGAYD IIICDECHST DATSILGIGT VLDQAETAGA
    HCV_J1 TYSTYGKFLA DGGCSGGAYD IIICDECHST DATSVLGIGT VLDQAETAGA
  HCV_J483 TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDQAETAGA
    HCV_J8 TYSTYGKFIA DGGCAAGAYD IIICDECHSV DATTILGIGT VLDQAETAGV
   HCV_JK1 TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTSILGIGT VLDQAETAGA
    HCV_JS TYSTYGKFLA DGGCSGGAYD IIMCDECHST DSTTILGIGT ALDQAETAGA
 HCV_K1_R1 TYSTYGKFLA DGGCSGGAYD IIMCDECHSI DSTSILGIGT VLDQAETAGA
 HCV_K1_R2 TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTSILGIGT VLDQAETAGA
 HCV_K1_R3 TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTSILGIGT VLDQAETAGA
 HCV_K1_S1 TYSTYGKFLA DGGCSGGAYD IIMCDECHSI DSTSILGIGT VLDQAETAGA
 HCV_K1_S2 TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTSILGIGT VLDQAETAGA
 HCV_K1_S3 TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTSILGIGT VLDQAETAGA
    HCV_L2 TYSTYGKFFA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDRAETAGA
     HCV_N TYSTYGKFLA DGGCSGGAYD IIMCDECHST DSTTILGIGT VLDQAETAGA
  HCV12083 TYSTYGEFLA DGGCSGGAYD IIICDECHST DPTTVLGVGT VLDQAETAGV
   HCV1480 TYSTYGKFFA DGGCSGGAYD VIICDECHSQ DATTILGIGT VLDQAETAGA
  HCVPOLYP TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDQAETAGA
      HD_1 TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTSILGIGT VLDQAETAGA
   HPCCGAA TYSTYGKFLA DAGCSGGAYD IIICDECHST DATSISGIGT VLDQAETAGA
     HPCFG TYSTYGKFLA DGGCSGGAYD VIICDECHAQ DATSILGIGT VLDQAETAGV
HPCGENANTI TYSTYGKFLA DGGCSGGAYD IIMCDECHST DSTTILGIGT VLDQAETAGA
  HPCGENOM TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDQAETAGA
  HPCHUMR TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDQAETAGA
     HPCJ TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDQAETAGA
   HPCJCG TYSTYCKFLA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDQAETAGA
  HPCJK046 TYSTYGKFLA DGGCSGGAYD VIICDECHST DPTTVLGIGT VLDQAETAGC
  HPCJK049 TYSTYGKFLA DGGCSGGAYD VIICDECHAQ DATSILGIGT VLDQAETAGA
   HPCJTA TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDQAETAGA
   HPCJTB TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTTILGIGT VLDQAETAGA
   HPCK3A TYSTYGKFLA GGGCSGGAYD VIICDDCHAQ DATSILGIGT VLDQAETAGV
 HPCPLYPRE TYSTYGKFLA DGGCSGGAYD IIICDECHST DATSILGIGT VLDQAETAGA
   HPCPOLP TYSTYGKFLA DGGCAGGAYD IIICDECHAV DSTTILGIGT VLDQAETAGV
     HPCPP TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTSILGIGT VLDQAETAGA
  HPCUNKCD TYSTYGKFLA DGGGSGGAYD IIMCDECHST DSTTIYGIGT VLDQAETAGA
   MKC1A TYSTYGKFLA DGGCSGGAYD IIICDECHST DSTSILGIGT VLDQAETAGA
   NDM59 TYSTYGKFLA DGGCAGGAYD IIICDECHAV DATTILGVGT VLDQAETAGV
    NZLI TYSTYGKFLA DGGCSGGAYD VIICDECHAQ DATSILGIGT VLDQAETAGV
    SA13 TYSTYGKFLA DGGCSGGAYD VIICDECHSQ DATTILGIGT VLDQAETAGA
    Th580 TYSTYGKFLA DGGCSGGAYD IIICDECHST DPTTVLGIGT VLDQAETAGV
Type_3a_CB TYSTYGKFLA DGGCSGGAYD VIICDECHAQ DATSILGIGT VLDQAETAGV
  TypeV_D TYSTYGKFLA DGGCSGGAYD VIICDECHAQ DATSILGIGT VLDQAETAGV
    VN004 TYSTYGKFLA DGGCSGGAYD VIICDECHST DPTTVSGIGT VLDQAETSGV
```

```
VN235   TYSTYGKFLA DGGCSGGAYD IIICDECHST DPTTVLGIGT VLDQAETAGV
VN405   TYSTYGKFLA DGGCSGGAYD IIICDECHST DPTTVLGIGT VLDQAETAGV

        1351                                                1400
BEBE1   RLTVLATATP PGSVTTPHPN IEEVALGHEG EIPFYGKAIP LSAIKGGRHL
D89815  RFVVLATATP PGSITFPHPN IEEVPLANTG EIPFYAKTIP IEVIRGGRHL
ED43type_4 RLTVLATATP PGSVTTPHSN IEEVALPTTG EIPFYGKAIP LELIKGGRHL
HC_C2   RLVVLATATP PGSVTVPHSN IEEVALSTTG EIPFYGKAIP IETIKGGRHL
HC_G9   RLTILATATP PGSVTVPHSN IEEVALSTEG EIPFYGKAIP LNYIKGGRHL
HCU16326 RLVVLSTATP PGSVTVPHLN IEEVALSNTG EIPFYGKAIP IEAIKGGRHL
HCV_H_CMR RLVVLATATP PGSVTVSHPN IEEVALSTTG EIPFYGKAIP LEVIKGGRHL
HCV_J1  RLVVLATATP PGSITVPHAN IEEVALSTTG EIPFYGKAIP LEAIKGGRHL
HCV_J483 RLVVLATATP PGSVTVPHPN IEEIGLSNNG EIPFYGKAIP IEAIKGGRHL
HCV_J8  RLVVLATATP PGTVTTPHSN IEEVALGHEG EIPFYGKAIP LAFIKGGRHL
HCV_JK1 RLVVLAAATP PGSVTVPHPN IEEVALPNTG EIPFYGKAIP LETIKGGRHL
HCV_JS  RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP IEVIKGGRHL
HCV_K1_R1 RLVVLATATP PGSVTVPHPN IEEIALSNTG EIPFYGKAIP IETIKGGRHL
HCV_K1_R2 RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP IETIKGGRHL
HCV_K1_R3 RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP LDTIKGGRHL
HCV_K1_S1 RLVVLATATP PGSVTVPHPN IEEIALSNTG EIPFYGKAIP IETIKGGRHL
HCV_K1_S2 RLVVLATATP PGSVTVPHPN IEEVALSNIG EIPFYGKAIP IETIKGGRHL
HCV_K1_S3 RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP LDTIKGGRHL
HCV_L2  RLVVLATATP PGSTTVPHPN IEEVALPNTG EIPFYGRAIP IEFIKGGRHL
HCV_N   RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP IEAIMGGRHL
HCV12083 RLTVLPTATP PGSVTVPHPN ITETALPTTG EIPFYGKAIP LEYIKGGRHL
HCV1480 RLVVLATAIP PGSVTTPHPN IEEVALPSEG EIPFYGRAIP LVLIKGGRHL
HCVPOLYP RLVVLATATP PGSVTVPHPN IEEVALSNIG EIPFYGKAIP IETIKGGRHL
HD_1    RLVVLATATP PGSVTVPHSN IEEVALSNIG EIPFYGKAIP LENIKGGRHL
HPCCGAA RLVVLATATP PGSVTVSHPN IEEVALSTTG EIPFYGKAIP LEVIKGGRHL
HPCFG   RLTVLATATP PGSITVPHPN IEEVGLTSDG EIPFYGKALP LAMIKGGRHL
HPCGENANTI RLVVLATATP PGSVTVPHPN IEEIALSNTG EIPFYGKAIP IETIKGGRHL
HPCGENOM RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP IEAIRGGRHL
HPCHUMR RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP IEAIRGGRHL
HPCJ    RLVVLATATP PGSVTVPHPN IEEVGLSNTG EIPFYGKAIP IEVIKGGRHL
HPCJCG  RLVVLATATP PGSITVPHPN IEEVALSNTG EIPFYGKAIP IEAIKGGRHL
HPCJK046 RLTVLATATP PGSVTVPHPN IQETALPLTG EVPFYGKAIP LEYIKGGRHL
HPCJK049 RLVVLATATP PGSITVPHSN IEEVALTGEG EIPFYGRAIP LGVIKGGRHL
HPCJTA  RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP LEAIKGGRHL
HPCJTB  RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP IEVIKGGRHL
HPCK3A  RLTVLATATP PGSITVPHSN IEEVALGSEG EIPFYGKAIP IACIKGGRHL
HPCPLYPRE RLVVLATATP PGSVTVPHPN IEEVALSTTG EIPFYGKAIP LEVIKGGRHL
HPCPOLP RLTVLATATP PGSVTTPHPN IEEVALGQEG EIPFYGRAIP LSYIKGGRHL
HPCPP   RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP LEAIKGGRHL
HPCUNKCD RLVVLSTATP PGSVTVPHLN IEEVALSNTG EIPFYGKAIP IEAIKGGRHL
MKC1A   RLVVLATATP PGSVTVPHPN IEEVALSNTG EIPFYGKAIP LEAIKGGRHL
NDM59   RLTVLATATP PGSVTTPHPN IEEVALGQEG EIPFYGRAIP LSYIKGGRHL
NZLI    RLTVLATATP PGSITVPHSN IEEVALGSEG EIPFYGKAIP IALLKGGRHL
SA13    RLVVLATATP PGSVTTPHPN IEEVALPSEG EIPFYGRAIP LALIKGGRHL
Th580   RLTVLATATP PGSVTVPHPN ITETALPTTG EIPFYGKCIP LEFIKGGRHL
Type_3a_CB RLTVLATRTP PGSITVPHSN IEEVALGSEG EIPFYGKAIP IALLKGGRHL
TypeV_D RLTVLATATP PGSITVPHSN IEEVALGSEG EIPFYGKAIP IALLKGGRHL
VN004   RLTVLATATP PGSVTVPHPN ITESALPTTG EIPFYGKAVP LEYIKGGRHL
VN235   RLTVLATATP PGSVTVPHPN ITETALPSTG EVPFYGKAIP LECIKGGRHL
VN405   RLTVLATATP PGSVTVPHPN ITEVALSSTG EVPFYGKAIP LEYIKGGRHL

        1401                                                1450
BEBE1   IFCHSKKKCD ELAVALRGMG LNAVAYYRGL DVSIIPTQGD VVVVATDALM
D89815  IFCHSKKKCD ELPAKLSALG LNAVAYYRGL DVSVIPASGD VVVVATDALM
ED43type_4 IFCHSKKKCD ELARQLTSLG LNAVAYYRGL DVSVIPTSGD VVVCATDALM
HC_C2   IFCHSKKKCD ELAAKLSALG LNAVAYYRGL DVSVIPTSGD VVVVATDALM
HC_G9   IFCHSKKKCD ELAAKLVGLG VNAVAFYRGL DVSVIPTTGD VVVVATDALM
```

```
HCU16326  IFCHSKKKCD ELAAKLSGLG LNAVAYYRGL DVSVIPTSGD VVVVATDALM
HCV_H_CMR IFCHSKKKCD ELAAKLVALG INAVAYYRGL DVSVIPTNGD VVVVSTDALM
  HCV_J1  IFCHSKKKCD ELAAKLVALG VNAVAYYRGL DVSVIPTSGD VVVVATDALM
 HCV_J483 IFCHSKKKCD ELAAKLTGLG LNAVAYYRGL DVSVIPPIGD VAVVATDALM
  HCV_J8  IFCHSKKKCD ELAAALRGMG VNAVAYYRGL DVSVIPTQGD VVVVATDALM
 HCV_JK1  IFCHSKKKCD ELAAKLSALG VNAVAYYRGL DVSVIPTSGD VVVVATDALM
  HCV_JS  IFCHSKKKCD ELAAKLSTLG LNAVAYYRGL DVSVIPTSGD VVVVATDALM
HCV_K1_R1 IFCHSKKKCD ELAAKLSGLG LNAVAYYRGL DVSVIPASGD VVVVATDALM
HCV_K1_R2 IFCHSKKKCD ELAAKLSGLG INAVAYYRGL DVSVIPASGD VVVVATDALM
HCV_K1_R3 IFWHSKKKCD ELATKLSALG VNAVAYYRGL DVSVIPTSGN VVVVATDALM
HCV_K1_S1 IFCHSKKKCD ELAAKLSGLG LNAVAYYRGL DVSVIPASGD VVVVATDALM
HCV_K1_S2 IFCHSKKKCD ELAAKLSGLG INAVAYYRGL DVSVIPASGD VVVVATDALM
HCV_K1_S3 IFWHSKKKCD ELATKLSALG VNAVAYYRGL DVSVIPTSGN VVVVATDALM
  HCV_L2  IFCPSKKKCD ELAAKLSALG INAVAYYRGL DVSVIPTSGD VVVVATDALM
   HCV_N  IFCHSKKKCD ELAAKLSGLG INAVAYYRGL DVSVIPTSGD VVVVATDALM
 HCV12083 IFCHSKKKCD ELAGKLKSLG LNAVAFYRGV DVSVIPTSGD VVVCATDALM
 HCV1480  IFCHSKKKCD ELAKQLTSLG VNAVAYYRGL DVAVIPATGD VVVCSTDALM
HCVPOLYP  IFCHSKKKCD ELAAKLSGLG LNAVAYYRGL DVSVIPTSGD VVVVATDALM
   HD_1   IFCHSRKKCD ELAAKLSGLG INAVAYYRGL DVSVIPTSGD VVVVATDALM
 HPCCGAA  IFCHSKKKCD ELAAKLVALG INAVAYYRGL DVSVIPTSGD VVVVATDALM
  HPCFG   VFCHSKEKCD ELASKLRGMG VNAVAFYRGL DVSVIPVSGD VVVCATDALM
HPCGENANTI IFCHSKKKCD ELAAKLSALG IHAVAYYRGL DVSVIPASGN VVVVATDALM
 HPCGENOM IFCHSKKKCD ELAAKLSSLG LNAVAYYRGL DVSVIPSSGD VVVVATDALM
 HPCHUMR  IFCHSKKKCD ELAAKLSGLG INAVAYYRGL DVSVIPTIGD VVVVATDALM
  HPCJ    IFCHSKKKCD ELAAKLSALG INAVAYYRGL DVSVIPASGD VVVVATDALM
 HPCJCG   IFCHSKKKCD ELAAKLTGLG LNAVAYYRGL DVSVIPTSGD VVVVATDALM
HPCJK046  IFCHSKKKCD ELAAQLRTLG LNAVAFYRGV DVSVIPTSGD VVVCATDALM
HPCJK049  IFCHSKKKCD ELAKQLTSLG VNAVAFYRGL DVSVIPTQGD VVVCATDALI
 HPCJTA   IFCHSKKKCD ELAAKLSGLG INAVAYYRGL DVSVIPTSGD VVIVATDALM
 HPCJTB   IFCHSKKKCD ELAAKLSGLG INAVAYYRGL DVSVIPTSGD VVVVATDALM
 HPCK3A   IFCHSKKKCD KMASKLRGMG LNAVAYYRGL DVSVIPTTGD VVVCATDALM
HPCPLYPRE IFCHSKKKCD ELAAKLVALG INAVAYYRGL DVSVIPTSGD VVVVATDALM
 HPCPOLP  IFCHSKKKCD ELAAALRGMG LNAVAYYRGL DVSVIPTQGD VVVVATDALM
  HPCPP   IFCHSKKKCD ELAAKLSALG VNAVAYYRGL DVSIIPTSGD VVVVATDALM
HPCUNKCD  IFCHSKKKCD ELAAKLSGLG LNAVAYYRGL DVSVIPTSGD VVVVATDALM
  MKC1A   IFCHSKKKCD ELAAKLSALG VNAVAYYRGL DVSIIPTSGD VVVVATDALM
  NDM59   IFCHSKKKCD ELSAALRSMG LNAVAYYRGL DVSVIPTQGD VVVVATDALM
  NZLI    IFCHSKKKCD EIASKLRGMG LNAVAYYRGL DVSVIPTTGD VVVCATDALM
  SA13    IFCHSKKKCD ELAKQLTSQG VNAVAYYRGL DVAVIPATGD VVVCSTDALM
  Th580   IFCHSKKKCD ELSKQLTSLG LNAVAFYRGV DVAVIPTSGD VVVCATDALM
Type_3a_CB IFCHSKKKCD EIASKLRGMG LNAVAYYRGL DVSVIPTTGD VVVCATDALM
 TypeV_D  IFCHSKKKCD EIASKLRGMG LNAVAYYRGL DVSVIPTTGD VVVCATDALM
  VN004   IFCHPKKKCD ELAKQLVSLG LNAVAFYRGV DVSVIPTSGD VVVCATDALM
  VN235   IFCHSKKKCD ELAKQLRTLG LNAVAFYRGV DVSVIPTAGD VVVCATDALM
  VN405   IFCHSKKKCD ELAKQLTSLG LNAVAFYRGV DVSVIPTSGD VVVCATDALM

          1451                                                1500
  BEBE1   TGYTGDFDSV IDCNVAVTQV VDFSLDPTFT ITTQTVPQDS VSRSQRRGRT
  D89815  TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRTQRRGRT
ED43type_4 TGFTGDFDSV IDCNTSVIQT VDFSLDPTFS IEITTVPQDA VSRSQRRGRT
  HC_C2   TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
  HC_G9   TGYTGDFDSV IDCNTCVVQT VDFSLDPTFS IETSTVPQDA VSRSQRRGRT
HCU16326  TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
HCV_H_CMR TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTLPQDA VSRTQRRGRT
  HCV_J1  TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTLPQDA VSRTQRRGRT
 HCV_J483 TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
  HCV_J8  TGYTGDFDSV IDCNVAVSQI VDFSLDPTFT ITTQTVPQDA VSRSQRRGRT
 HCV_JK1  TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTLPQDA VSRSQRRGRT
  HCV_JS  TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
HCV_K1_R1 TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
HCV_K1_R2 TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
```

```
HCV_K1_R3   TGYTGDFDSV IDCNTCVIQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
HCV_K1_S1   TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
HCV_K1_S2   TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
HCV_K1_S3   TGYTGDFDSV IDCNTCVIQT DDFSLDPTFT IETRTVPQDA VSRSQRRGRT
   HCV_L2   TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRTQRRGRT
    HCV_N   TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
  HCV12083  TGYTGDFDSV IDCNVAVTQV VDFSLDPTFS IETTTVPQDA VSRSQRRGRT
  HCV1480   TGFTGDFDSV IDCNSAVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
  HCVPOLYP  TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
     HD_1   TGYTGDFDSV IDCNVCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
  HPCCGAA   TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTLPQDA VSRTQRRGRT
    HPCFG   TGYTGDFDTV IDCNVAVEQY VDFSLDPTFS IETRTVPQDA VSRSQRRGRT
HPCGENANTI  TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTMPQDA VSRSQRRGRT
  HPCGENOM  TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
  HPCHUMR   TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
    HPCJ    TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
   HPCJCG   TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTLPQDA VSRAQRRGRT
  HPCJK046  TGYTGDFDSV IDCNVAVTQI VDFSLDPTFS IETTTVPQDA VARSQRRGRT
  HPCJK049  TGYTGDFDSV IDCNVAVEQY VDFSLDPTFS IETHTVPQDA VSRSQRRGRT
   HPCJTA   TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
   HPCJTB   TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
   HPCK3A   TGFTGDFDSV IDCNVAVEQY VDFSLDPTFS IETCTAPQDA VSRSQRRGRT
 HPCPLYPRE  TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETITLPQDA VSRTQRRGRT
   HPCPOLP  TGFTGDFDSV IDCNVAVTQV VDFSLDPTFT ITTQTVPQDA VSRSQRRGRT
    HPCPP   TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
  HPCUNKCD  TGFTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
    MKC1A   TGYTGDFDSV IDCNTCVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
    NDM59   TGYTGDFDSV IDCNVAVTQV VDFSLDPTFT ITTQTVPQDA VSRSQRRGRT
     NZLI   TGFTGDFDSV IDCNVAVEQY VDFSLDPTFS IETRTAPQDA VSRSQRRGRT
     SA13   TGFTGDFDSV IDCNTTVTQT VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
    Th580   TGYTGDFDSV IDCNVAVTQV VDFSLDPTFS IETTTVPQDA VSRSQRRGRT
Type_3a_CB  TGYTGDFDSV IDCNVAVEQY VDFSLDPTFS IETRTAPQDA VSRSQRRGRT
  TypeV_D   TGFTGDFDSV IDCNVAVEQY VDFSLDPTFS IETRTAPQDA VSRSQRRGRT
    VN004   TGYTGDFDSV IDCNVTVTQV VDFSLDPTFT IETTTVPQDA VSRSQRRGRT
    VN235   TGYTGDFDSV IDCNVAVTQI VDFSLDPTFS IETTTVPQDA VARSQRRGRT
    VN405   TGYTGDFDSV IDCNVSVTQV VDFSLDPTFT IETTTMPQDA VSRSQRRGRT

            1501                                                1550
    BEBE1   GRGRLGIYRY VSSGERASGM FDTVVLCECY DAGAAWYELT PAETTVRLRA
    D89815  GRGRRGIYRF VTPGERPSAM FDSSVLCECY DAGCAWYELT PAETSVRLRA
ED43type_4  GRGRLGTYRY VTPGERPSGM FDTAELCECY DAGCAWYELT PAETTTRLKA
     HC_C2  GRGRRGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
     HC_G9  GRGKHGIYRY VSPGERPSGM FDSVVLCECY DAGCAWYELT PAETTVRLRA
  HCU16326  GRGRAGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
  HCV_H_CMR GRGKPGIYRF VAPGERPSGM FDSSVLCECY DAGCAWYELM PAETTVRLRA
    HCV_J1  GRGKPGIYRF VAPGERPSGM FDSSILCECY DTGCAWYELT PAETTVRLRA
  HCV_J483  GRGRSGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
    HCV_J8  GRGRLGVYRY VSSGERPSGM FDSVVLCECY DAGAAWYELT PAETTVRLRA
   HCV_JK1  GRGRGGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAVTSVRLRA
    HCV_JS  GRGRGGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
  HCV_K1_R1 GRGRRGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
  HCV_K1_R2 GRGRAGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
  HCV_K1_R3 GRGRRGIYRF VTPGERTSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
  HCV_K1_S1 GRGRRGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
  HCV_K1_S2 GRGRAGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
  HCV_K1_S3 GRGRRGIYRF VTPGERTSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
    HCV_L2  GRGRGGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETTVRLRA
     HCV_N  GRGRGGIYRF VTPGERPSGM FDSPVLCECY DAGCAWYELT PAETSVRLRA
  HCV12083  GRGKPGVYRF VSQGERPSGM FDTVVLCEAY DTGCAWYELT PSETTVRLRA
  HCV1480   GRGRHGIYRY VSSGERPSGI FDSVVLCECY DAGCAWYDLT PAETTVRLRA
  HCVPOLYP  GRGRRGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
```

```
HD_1          GRGRMGIYRF VTPGERPSGM FDSSVLCESY DAGCAWYELT PAEASVRLRA
HPCCGAA       GRGKPGIYRF VAPGERPSGM FDSSVLCECY DAGCAWYELT PAETTVRLRA
HPCFG         GRGRPGIYRF VTPGERPSGM FDSVVLCECY DAGCSWYDLQ PAETTVRLRA
HPCGENANTI    SRGRRGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
HPCGENOM      GRGREGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETTVRLRA
HPCHUMR       GRGRRGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
HPCJ          GRGRAGIYRF VTPGERPSGM FDSSVLCECY DSGCAWYELT PAETSVRLRA
HPCJCG        GRGRSGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
HPCJK046      GRGKPGVYRY VSQGERPSGM FDTVVLCEAY DTGAAWYELT PAETTVRLRA
HPCJK049      GRGKSGTYRY VSPGERPSGM FDSVVLCECY DAGCAWYELT PSETTVRLRA
HPCJTA        GRGRGGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETTVRLRA
HPCJTB        GRGRGGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETTVRLRA
HPCK3A        GRGRLGTYRY VTPGERPSGM FDSVVLCECY DAGCSWYDLQ PAETTVRLRA
HPCPLYPRE     GRGKPGIYRF VAPGERPSGM FDSSVLCECY DAGCAWYELT PAETTVRLRA
HPCPOLP       GRGRLGIYRY VSTGERASGM FDSVVLCECY DAGAAWYELT PAETTVRLRA
HPCPP         GRGRGGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
HPCUNKCD      GRGRAGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
MKC1A         GRGRGGIYRF VTPGERPSGM FDSSVLCECY DAGCAWYELT PAETSVRLRA
NDM59         GRGRLGIYRY VSTGERASGM FDSVVLCECY DAGAAWYELT PSETTVRLRA
NZLI          GRGRLGTYRY VASGERPSGM FDSVVLCECY DAGCSWYDLQ PAETTVRLRA
SA13          GRGRHGIYRY VSSGERPSGI FDSVVLCECY DAGCAWYDLT PAETTVRLRA
Th580         GRGKPGVYRF VSQGERPSGM FDSVVLCEAY DTGCAWYELT PAETTVRLRA
Type_3a_CB    GRGRLGTYRY VAPGERPSGM FDSVVLCECY DAGCSWYDLQ PAETTVRLRA
TypeV_D       GRGRLGTYRY VAPGERPSGM FDSVVLCECY DAGCSWYDLQ PAETTVRLRA
VN004         GRGKHGVYRY VSQGERPSGM FDSVILCEAY DTGCAWYELT PAETTVRLRA
VN235         GRGKPGVYRY VSQGERPSGM FDTVVLCEAY DVGCAWYELT PSETTVRLRA
VN405         GRGKHGVYRY VSQGERPSGI FDTVVLCEAY DTGCAWYELT PSETTVRLRA

              1551                                                1600
BEBE1         YFNTPGLPVC QDHLEFWEAV FTGLTHIDAH FLSQTKQAGE GFPYLVAYQA
D89815        YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
ED43type_4    YFDTPGLPVC QDHLEFWESV FTGLTHIDGH FLSQTKQSGE NFPYLVAYQA
HC_C2         YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HC_G9         YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQSGE NFPYLVAYQA
HCU16326      YLNTPGLPVC QDHLEFSEGV FTGLTHIDAH FLSQTKQAGE NFPYLVAYQA
HCV_H_CMR     YMNTPGLPVC QDHLEFWEGV FTGLTHIDAH FLSQTKQSGE NFPYLVAYQA
HCV_J1        YMNTPGLPVC QDHLEFWEGV FTGLTHIDAH FLSQTKQGGE NFPYLVAYQA
HCV_J483      YLNTPGLPVC QDHLEFWESV FTGLSHIDAH FLSQTKQAGD NFPYLVAYQA
HCV_J8        YFNTPGLPVC QDHLEFWEAV FTGLTHIDAH FLSQTKQGGE NFAYLTAYQA
HCV_JK1       YLNTPGLPVC QVHLEFWESV FTGLTHIDAH FLSQTKQAGE NFPYLVAYQA
HCV_JS        YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HCV_K1_R1     YLNTPGLPFC QDHLEFWEGV FTGLTHIDAH FLSQTKQAGE NLPYLVAYQA
HCV_K1_R2     YLNTPGLPVW QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HCV_K1_R3     YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HCV_K1_S1     YLNTPGLPFC QDHLEFWEGV FTGLTHIDAH FLSQTKQAGE NLPYLVAYQA
HCV_K1_S2     YLNTPGLPVW QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HCV_K1_S3     YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HCV_L2        YLNTPGLPVC QDHLEFWESV FTGLNHIDAH FLSQTKQAGD NFPYLVAYQA
HCV_N         YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HCV12083      YMNTPGLPVC QDHLEFWEGV FTGLTHIDAH FLSHTKQAGE NFAYLVAYQA
HCV1480       YLNTPGLPVC QEHLEFWEGV FTGLTNIDAH MLSQAKQGGE NFPYLVAYQA
HCVPOLYP      YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGE NFPYLTAYQA
HD_1          YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HPCCGAA       YMNTPGLPVC QDHLGFWEGV FTGLTHIDAH FLSQTKQSGE NFPYLVAYQA
HPCFG         YLSTPGLPVC QDHLDFWERV FTGLTHIDAH FLSQAKQQGL NFAYLVAYQA
HPCGENANTI    YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HPCGENOM      YLNTPGLPVC QDHLEFWEGV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HPCHUMR       YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HPCJ          YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
HPCJCG        YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NLPYLVAYQA
HPCJK046      YLNTPGLPVC QDHLEFWEAV FTGLTHIDAH FLSQTKQGGE NFAYLVAYQA
```

```
HPCJK049   YLSTPGLPVC QDHLEFWEGV FTGLTHIDAH FLSQTKQQGL NFPYLTAYQA
  HPCJTA   YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
  HPCJTB   YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVVYQA
  HPCK3A   YLSTPGLPVC QDHLDLWESV FTGLTHIDAH FLSQTKQAGL NFSYLTAYQA
HPCPLYPRE  YMNTPGLPVC QDHLEFWEGV FTGLTHIDAH FLSQTKQSGE NLPYLVAYQA
 HPCPOLP   YFNTPGLPVC QDHLEFWEAV FTGLTHIDAH FLSQTKQSGE NFAYLTAYQA
   HPCPP   YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
 HPCUNKCD  YLNTPGLPVC QDHLEFSEGV FTGLTHIDAH FLSQTKQAGE NFPYLVAYQA
   MKC1A   YLNTPGLPVC QDHLEFWESV FTGLTHIDAH FLSQTKQAGD NFPYLVAYQA
   NDM59   YFNTPGLPVC QDHLEFWEAV FTGLTHIDAH FLSQTKQSGE NFAYLVAYQA
    NZLI   YLSTPGLPVC QDHLDFWESV FTGLTHIDAH FLSQTKQQGL NFSYLTAYQA
    SA13   YLNTPGLPVC QDHLEFWEGV FTGLTNIDAH MLSQTKQGGE NFPYLVAYQA
   Th580   YLNTPGLPVC QDHLEFWEGV FTGLTHIDAH FLSQTKQGGE NFAYLVAYQA
Type_3a_CB YLSTPGLPVC QDHLDFWESV FTGLTHIDAH FLSQTKQQGL NFSYLTAYQA
   TypeV_D YLSTPGLPVC QDHLDFWESV FTGLTHIDAH FLSQTKQQGL NFSYLTAYQA
    VN004  YLNTPGLPVC QDHLEFWEGV FTGLTHIDAH FLSQTKQAEE NFAYLVAYQA
    VN235  YLNTPGLPVC QDHLEFWEGV FTGMTHIDAH FLSQTKQGGE NFAYLVAYQA
    VN405  YLNTPGLPVC QDHLEFWEGV FTGLTHIDAH LLSQTKQGGE NFAYLVAYQA

           1601                                                1650
    BEBE1  TVCARAKAPP PSWDVMWKCL IRLKPTLVGP TPLLYRLGSV TNEVTLTHPV
   D89815  TVCARAKAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
ED43type_4 TVSAKVWLAP PSWDTMWKCL IRLKPTLHGP TPLLYRLGSV QNEVVLTHPI
   HC_C2   TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
   HC_G9   TVCARAKAPP PSWDQMWKCL IRLKPTLTGA TPLLYRLGGV QNEITLTHPI
 HCU16326  TVCARAQAPP PSWDEMWRCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
 HCV_H_CMR TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
   HCV_J1  TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
  HCV_J483 TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QGEVTLTHPV
   HCV_J8  TVCARAKAPP PSWDVMWKCL TRLKPTLTGP TPLLYRLGAV TNEVTLTHPV
   HCV_JK1 TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
   HCV_JS  TVCARAQASP PSWDQMWKCL TRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
 HCV_K1_R1 TVCARAQAPP PSWDQMWKCL TRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
 HCV_K1_R2 TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNDVTLTHPI
 HCV_K1_R3 TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVILTHPI
 HCV_K1_S1 TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
 HCV_K1_S2 TVCARAQAPP PSWDQMWKCL TRLKPTLHGP TPLLYRLGAV QNDVTLTHPI
 HCV_K1_S3 TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVILTHPI
   HCV_L2  TVCARAQAPP PSWDQMWKCL TRLKPTLHGP TPLLYRLGAV QNEVILTHPI
   HCV_N   TVCARAQAPP PSWDQMWKCL IWLKPVLHGP TPLLYRLGAV QNEITLTHPI
  HCV12083 TVCARAKAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
  HCV1480  TVCARAKAPP PSWDMMWKCL IRLKPTLTGP TPLLYRLGAV QNGVITTHPI
  HCVPOLYP TVCVRAKAPP PSWDTMWKCM ICLKPTLTGP TPLLYRLGAV QNEITLTHPI
   HD_1    TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVVLTHPI
  HPCCGAA  TVCARAQAPP PSWDQMRKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
   HPCFG   TVCARAKASP PCWDEMWKCL IRLKPTLQGP TPLLYRLGAI QNDICMTHPI
HPCGENANTI TVCARAQAPP PSWDQMWKCL TRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
  HPCGENOM TVCAKAQAPP PSWDQMWKCL TRLKPTLQGP TPLLYRLGAV QNEVTLTHPI
  HPCHUMR  TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
   HPCJ    TVCARSQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
  HPCJCG   TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
  HPCJK046 TVCARAKAPP PSWDTMWKCL LRLKPTLTGP TPLLYRLGAV QNEVTPTHPV
  HPCJK049 TVCARAAALP PSWDETWKCL IRLKPTLHGP TPLLYRLGAV QNEICTTHPV
  HPCJTA   TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEITLTHPI
  HPCJTB   TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEITLTHPI
  HPCK3A   TVCARAQAPP PSWDETWKCL VRLKPTLHGP TPLLYRLGPV QNEICLTHPI
HPCPLYPRE  TVCARAQAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEITLTHPV
  HPCPOLP  TVCARAKAPP PSWDVMWKCL TRLKPTLVGP TPLLYRLGSV TNEVTLTHPV
   HPCPP   TVCARAKAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
 HPCUNKCD  TVCARAQAPP PSWDEMWRCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
   MKC1A   TVCARAKAPP PSWDQMWKCL IRLKPTLHGP TPLLYRLGAV QNEVTLTHPI
```

```
      NDM59  TVCARAKAPP PSWDVMWKCL TRLKPTLVGP TPLLYRLGPV TNEVTLTHPV
       NZLI  TVCARAQAPP PSWDEMWKCL VRLKPTLHGP TPLLYRLGPV QNETCLTHPI
       SA13  TVCVRAKAPP PSWDTMWKCM LRLKPTLTGP TPLLYRLGAV QNEITLTHPI
       Th580 TVCARAKAPP PSWDVMWKCL TRLKPTLTGP TPLLYRLGAV QNEITLTHPI
   Type_3a_CB TVCARAQAPP PSWDETWKCL VRLKPTLHGP TPLLYRLGAV QNEIVTTHPI
     TypeV_D TVCARAQAPP PSWDEMWKCL VRLKPTLHGP TPLLYRLGPV QNEICLSHPI
       VN004 TVCARAKAPP PSWDTMWKCL IRLKPMLTGP TPLLYRLGPV QNETCLTHPV
       VN235 TVCARAKAPP PSWDTMWKCL IRLKPMLTGP TPLLYRLGPV QNEVVTTHPI
       VN405 TVCARAKAPP PSWDTMWKCL IRLKPMLTGP TPLLYRLGAV QNEIITTHPI
       VN405 TVCARAKAPP PSWDTMWKCL IRLKPMLTGP TPLLYRLGAV QNEITTTHPI

             1651                                                1700
       BEBE1 TKYIATCMQA DLEIMTSTWV LAGGVLAAVA AYCLATGCVS IIGRIHVNQK
       D89815 TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
   ED43type_4 TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLSVGSVV IVGRVVLSGQ
       HC_C2 TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIVLSGR
       HC_G9 TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLSTGSVV IVGRIILSGK
     HCU16326 TKFIMTCMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
     HCV_H_CMR TKYIMTCMSA DLEVVTSTWV LVGGVLAALA AYCLSTGCVV IVGRIVLSGK
       HCV_J1 TKYIMTCMSA DLEVVTSTWV LVGGVLAALA AYCLSTGCVV IVGRIVLSGR
     HCV_J483 TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
       HCV_J8 TKYIATCMQA DLEIMTSSWV LAGGVLAAVA AYCLATGCIS IIGRLHLNDR
     HCV_JK1 TKFIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
       HCV_JS TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
     HCV_K1_R1 TKYIMTCMSA DLEVVTSTWV LVGGVLAALT AYCLTTGSVV IVGRIILSGK
     HCV_K1_R2 TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
     HCV_K1_R3 TKFIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
     HCV_K1_S1 TKYIMTCMSA DLEVVTSTWV LVGGVLAALT AYCLTTGSVV IVGRIILSGK
     HCV_K1_S2 TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
     HCV_K1_S3 TKFIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
       HCV_L2 TKLIMASMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
       HCV_N TKYIMACMWA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
     HCV12083 TKYIMTCMSA DLEVITSTWV LVGGVLAALA AYCLSVGCVV ICGRITLTGK
     HCV1480 TKYIMACMSA DLEVITSTWV LVGGVVAALA AYCLTVGSVA IVGRIILSGR
     HCVPOLYP TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
       HD_1 TKFIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGCVV IVGRIILSGR
     HPCCGAA TKYIMTCMSA DLEVVTSTWV LVGGVLAALA AYCLSTGCVV IVGRIVLSGK
       HPCFG TKYIMACMSA DLEVTTSAWV LVGGVLAALA AYCLSVGCVV IVGHIELGGK
   HPCGENANTI TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
     HPCGENOM TKYIMTCMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIVLSGS
     HPCHUMR TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
       HPCJ TKFIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
     HPCJCG TKYIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
     HPCJK046 TKYIMACMSA DLEVITSTWV VAGGILAAIA AYCLTVGSVV ICGRITTSSR
     HPCJK049 TKYIATCMAA DLEVATSAWV LLGGVMAALT AYCLSVGSVV IVGHLVLGGK
     HPCJTA TKFIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
     HPCJTB TKFIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
     HPCK3A TKYVMACMSA DLEVTTSTWV LLGGVLAAVA AYCLSVGCVV IVGHIELGGK
   HPCPLYPRE TKYIMTCMSA DLEVVTSTWV LVGGVLAALA AYCLSTGCVV IVGRVVLSGK
     HPCPOLP TKYIATCMQA DLEVMTSTWV LAGGVLAAVA AYCLATGCVC IIGRLHVNQR
     HPCPP TKFIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
     HPCUNKCD TKFIMTCMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGK
       MKC1A TKFIMACMSA DLEVVTSTWV LVGGVLAALA AYCLTTGSVV IVGRIILSGR
       NDM59 TKYIATCMQA DLEVMTSTWV LAGGVLAAVA AYCLATGCVS IIGRLHVNQR
       NZLI TKYLMACMSA DLEVTTSTWV LLGGVLAALA AYCLSVGCVV IVGHIELEGK
       SA13 TKYIMACMSA DLEVITSTWV LVGGVVAALA AYCLTVGSVA IVGRIILSGR
       Th580 TKYIMTCMSA DLEVITSTWV IVGGVLAALA AYCLTVGCVV ICGRIVTSGK
   Type_3a_CB TKYVMACMSA DLEVTTSTWV LLGGVLAALA AYCLSVGCVV IVGHIELGGK
     TypeV_D TKYIMACMSA DLEVTTSTWV LLGGVLAALA AYCLSVGCVV IVGHIELGGK
       VN004 TKYIMTCMSA DLEVITSTWV LVGGVVAALA AYCLSVGCVV ICGRISTSGK
       VN235 TKYIMTCMAA DLEVITSTWV LAGGIVAALA AYCLTVGSVV ICGRIVTSGK
       VN405 TKYIMTCMSA DLEVITSTWV LVGGVLAALA AYCLSVGCVV VCGRISTTGK
```

```
                1701                                                    1750
      BEBE1     TIIAPDKEVL YEAFDEMEEC ASRTALIEEG HRIAEMLKSK IQGLMQQASK
      D89815    PAVIPDREVL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
  ED43type_4    PAVIPDREVL YQQFDEMEEC SKHLPLVEHG LQLAEQFKQK ALGLLNFAGK
       HC_C2    PAVIPDREVL YQEFDEMEEC GSHLPYIEQG MQLAEQFKQK ALGLLQIATK
       HC_G9    PAVIPDREVL YREFDEMEEC AAHIPYLEQG MHLAEQFKQK ALGLLQTASK
     HCU16326   PAIIPDREVL YQEFDEMEEC ASHLPYFEQG MQLAEQFKQK ALGLLQTATK
     HCV_H_CMR  PAIIPDREVL YQEFDEMEEC SQHLPYIEQG MMLAEQFKQK ALGLLQTASR
       HCV_J1   PAIIPDREVL YREFDEMEEC SQHLPYIEQG MMLAEQFKQK ALGLLQTASR
     HCV_J483   PAVVPDREVL YQEFDEMEEC ASQLPYIEQG MQLAEQFKQK ALGLLQTASR
       HCV_J8   VVVAPDKEIL YEAFDEMEEC ASKAALIEEG QRMAEMLKSK IQGLLQQATR
      HCV_JK1   PAIIPDREVL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTASK
       HCV_JS   PAVIPDREVL YREFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
     HCV_K1_R1  PAVIPDREAL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATN
     HCV_K1_R2  PAIIPDREVL YREFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
     HCV_K1_R3  PAVIPDREVL YREFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
     HCV_K1_S1  PAVIPDREAL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATN
     HCV_K1_S2  PAIIPDREVL YREFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
     HCV_K1_S3  PAVIPDREVL YREFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
       HCV_L2   PAVIPDREVL YREFDEMEEC ASHLPYIEQG VQLAEQFKQK ALGLLQTATK
        HCV_N   PAVVPDREVL YREFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
     HCV12083   PAVVPDREIL YQQFDEMEEC SRHIPYLAEG QQIAEQFRQK VLGLLQASAK
      HCV1480   PAITPDREVL YQQFDEMEEC SASLPYVDEA RAIAGQFKEK VLGLIGTAGQ
     HCVPOLYP   PAIIPDREVL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
         HD_1   PAIVPDREVL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
     HPCCGAA    PAIIPDREVL YQEFDEMEEC SQHLPYIEQG MMLAEQFKQK ALGLLQTASR
       HPCFG    PALVPDRQVL YQQYDEMEEC SQSAPYIEQA QAIAQQFKDK VLGLLQRASQ
  HPCGENANTI    PAVVPDREVL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
    HPCGENOM    PAIVPDREVL YQDFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
     HPCHUMR    PAIVPDRELL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
        HPCJ    PAVIPDREVL YREFDEMEEC ASHLPYIEQG MLLAEQFKQK ALGLLQMATK
      HPCJCG    PAVIPDREVL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
     HPCJK046   PAVIPDREVM YQQYDEMEEC SRHLPYLVEG QQLAEQFKQN VLGLIQVTTK
     HPCJK049   PALVPDKEVL YQQYDEMEEC SRAAPYIEQA QGIAQQFKEK VIGLLQQADQ
      HPCJTA    PAVVPDREVL YREFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
      HPCJTB    PAVVPDREVL YREFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
      HPCK3A    PALVPDKEVL YQQYDEMEEC SQARPYIEQA QVIAHQFKEK VLGLLQRATQ
    HPCPLYPRE   PAIIPDREVL YREFDEMEEC SQHLPYIEQG MMLAEQFKQK ALGLLQTASR
     HPCPOLP    AVVAPDKEVL YEAFDEMEEC ASRAALIEEG QRIAEMLKSK IQGLLQQASK
       HPCPP    PAVIPDREVL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
    HPCUNKCD    PAIIPDREVL YQEFDEMEEC ASHLPYFEQG MQLAEQFKQK ALGLLQTATK
       MKC1A    PAVIPDREVL YQEFDEMEEC ASHLPYIEQG MQLAEQFKQK ALGLLQTATK
       NDM59    AVVAPDKEVL YEAFDEMEEC ASRAALIEEG QRIAEMLKSK IQGLLQQASK
        NZLI    PALVPDKEVL YQQYDEMEEC SQAAPYIEQA QVIAHQFKEK ILGLLQRATQ
        SA13    PAIIPDREVL YQQFDEMEEC SASLPYMDEA RAIAEQFKEK VLGLIGTAGQ
       Th580    PAVVPDREVL YQQFDEMEEC SKHIPYLVEG QQIAEQFKQK VLGLLQAGTK
    Type_3a_CB  PALVPDKEVL YQQYDEMEEC SQAAPYIEQA QAIAHQFKEK VLGLLQRATQ
      TypeV_D   PALVPDKEVL YQQYDEMEEC SQAAPYIEQA QAIAHQFKEK VLGLLQRATQ
        VN004   PVLIPDREVL YQQFDEMEEC SRHIPYLAEG HLIAEQFKQK VLGLIQSTSK
        VN235   PVPLPDREVL YRQFDEMEEC SRHIPYLAEG QQIAEQFKQK ILGLLQNTAK
        VN405   PVLIPDREVL YQQFDEMEEC SRHIPYLVEG QHLAEQFKQK VLGLIQTTTR

                1751                                                    1800
      BEBE1     QAQGVQPAVQ ATWPKLEQFW AKHMWNFISG IQYLAGLSTL PGNPAVASMM
      D89815    QAEAAAPVVE SKWRALETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
  ED43type_4    QAQEATPVIQ SNFAKLEQFW ANDMWNFISG IQYLAGLSTL PGNPAIASLM
       HC_C2    QAEAAAPVVE SKWRALETFW AKHMWNFISG VQYLAGLSTL PGNPAIASLM
       HC_G9    QAETITPAVH TNWQKLESFW AKHMWNFVSG IQYLAGLSTL PGNPAIASLM
     HCU16326   QAEAAAPVVE SKWRALETFW AKHMWNFISG IQYLAGLSTL PGNPAIRSPM
     HCV_H_CMR  HAEVITPAVQ TNWQKLEVFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
```

```
HCV_J1          QAEVIAPTVQ TNWQKLEAFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV_J483        QAEAAAPVVE SKWRALETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV_J8          QAQDIQPAIQ SSWPKLEQFW AKHMWNFISG IQYLAGLSTL PGNPAVASMM
HCV_JK1         QAEAAAPVVE SKWQALEAFW AKHMWNFISG IQYLAGLSTL PGNPAIVSLM
HCV_JS          QAEAAAPVME SKWRALETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV_K1_R1       QAEAAAPVVE SKWRALEAFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV_K1_R2       QAEAAAPVVE SKWQALETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV_K1_R3       QAEAAAPVVE SKWRTLEVFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV_K1_S1       QAEAAAPVVE SKWRALEAFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV_K1_S2       QAEAAAPVVE SKWQALETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV_K1_S3       QAEAAAPVVE SKWRTLEVFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV_L2          QAEAAAPVVE SKWRALETFW AKHMWNFISG IQYLAALSTL PGNPAIASLM
HCV_N           QAEAAAPVVG SKWRAFETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HCV12083        QAEELKPAVH SAWPRVEDFW RKHMWNFVSG IQYLAGLSTL PGNPAVASLM
HCV1480         KAETLKPAAT SMWSKAEQFW AKHMWNFVSG IQYLAGLSTL PGNPAVATLM
HCVPOLYP        QAEAAVPVVE SKWQALEAFW AKHMWNFISG IQYLAGLSTL PGNLAIASLM
HD_1            QAEAAAPVVE SKWRALEAFW AKHMRNFISG IQYLAGLSTL PGNPAIASLM
HPCCGAA         HAEVITPAVQ TNWQKLEVFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HPCFG           QEAEIRPIVQ SQWQKAEAFW QQHMWNFVSG IQYLAGLSTL PGNPAVASLM
HPCGENANTI      QAEAAAPVVE SKWRTLEAFW ANDMWNFISG IQYLAGLSTL PGNPAIASLM
HPCGENOM        QAEAAAPVVE SKWRALETFW EKHMWNFISG IQYLAGLSTL PGNPAMASLM
HPCHUMR         QAEAAAPVVE SKWRALETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HPCJ            QAEAAAPVVE TKWQALEVFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HPCJCG          QAEAAAPVVE SKWRALEVFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HPCJK046        QAEELKPAVH SAWPKLEQFW YKHMWNFISG IQYLAGLSTL PGNPAVAALM
HPCJK049        KAADIKPIAT PYWQKLETFW SKHMWNFVSG IQYLAGLSTL PGNPAIASLM
HPCJTA          QAEAAAPVVE SRWRALEAFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HPCJTB          QAEAAAPVVE SRWRALEAFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HPCK3A          QQAVIEPIVV SNWQKLEVLW HKHMWNFVSG IQYLAGLSTL PGNPAVASLM
HPCPLYPRE       QAEVIAPAVQ TNWQKLETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HPCPOLP         QAQDIQPAVQ ASWPKVEQFW AKHMWNFISG IQYLAGLSTL PGNPAVASMM
HPCPP           QAEAAAPVVE SKWRALETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
HPCUNKCD        QAEAAAPVVE SKWRALETFW AKHMWNFISG IQYLAGLSTL PGNPAIRSPM
MKC1A           QAEAAAPVVE SKWRALETFW AKHMWNFISG IQYLAGLSTL PGNPAIASLM
NDM59           QAQDIQPAVQ ASWPKVEQFW AKHMWNFISG IQYLAGLSTL PGNPAVASMM
NZLI            QQAVIEPIVT TNWQKLEAFW HKHMWNFVSG IQYLAGLSTL PGNPAVASLM
SA13            KAETLKPAAT SMWNRAEQFW AKHMWNFVSG IQYLAGLSTL PGNPAVATLM
Th580           HAEELKPAIH STWPRVEEFW RKHMWNFVSG IQYLAGLSTL PGSPAVASLM
Type_3a_CB      QQAVIEPIVA TNWQKLEAFW HKHMWNFVSG IQYLAGLSTL PGNPAVASLM
TypeV_D         QQAVIEPIVA TNWQKLEAFW HKHMWNFVSG IQYLAGLPTL PGNPAVASLM
VN004           QAEELKPAVH AAWPKLEQFW QKQLWNFVSG IQYLAGLSTL PGNPAIASLM
VN235           QAEDLKPAVQ SAWPKLEQFW QKHLWNFVSG VQYLAGLSTL PGNPAVASLM
VN405           QAEEIEPVVH SAWPKLEQFW QKHLWNFVSG IQYLAGLSTL PGNPAVASLM

                1801                                                1850
BEBE1           SFSAALTSPL STSTTILLNI MGGWLASQIA PPAGATGFVV SGLVGAAVGS
D89815          AFTASITSPL ATQYTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
ED43type_4      SFTAAVTSPL TTQQTLLFNI LGGWVASQIR DSDASTAFVV SGLAGAAVGS
HC_C2           AFTASVTSPL TTQSTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
HC_G9           SFTAAVTSPL TTQQTLLFNI LGGWVAAQLA APAAATAFVG AGITGAVIGS
HCU16326        AFTASITSPL TTQHTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGT
HCV_H_CMR       AFTAAVTSPL TTGQTLLFNI LGGWVAAQLA APGAATAFVG AGLAGAAIGS
HCV_J1          AFTAAVTSPL TTSQTLLFNI LGGWVAAQLA APGAATAFVG SGLAGAAVGS
HCV_J483        AFTASITSPL TTQNTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
HCV_J8          AFSAALTSPL PTSTTILLNI MGGWLASQIA PPAGATGFVV SGLVGAAVGS
HCV_JK1         AFTASITSPL TTQHTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
HCV_JS          AFTASITSPL TTQSTLLFNI LGGWVAAQLA PPSAASAFVG AGIVGAAVGS
HCV_K1_R1       AFTASITSPL TTQSTLLFNI LGGWVAAQLA PPRAVSAFVG AGIAGAAVGS
HCV_K1_R2       AFTASITSPL TTQHTLLFNI LGGWVAAQLA PPRAASAFVG AGIAGAAVGS
HCV_K1_R3       AFTASITSPL TTQHTLLFNI LGGWVAAQLA PPRAASAFVG AGIAGAAVGS
HCV_K1_S1       AFTASITSPL TTQSTLLFNI LGGWVAAQLA PPRAVSAFVG AGIAGAAVGS
```

```
HCV_K1_S2   AFTASITSPL TTQHTLLFNI LGGWVAAQLA PPRAASAFVG AGIAGAAVGS
HCV_K1_S3   AFTASITSPL TTQHTLLFNI LGGWVAAQLA PPRAASAFVG AGIAGAAVGS
  HCV_L2    AFTASITSPL TTQNTLLFNI LGGWVAAQLA PASAASAFVG AGIAGAAVGS
   HCV_N    AFTASITSPL TTQNTLLFNI LGGWVAAQLA PPSAASAFVG AGSAGAAIGT
 HCV12083   SFTASLTSPL RTSQTLLLNI LGGWIAAQVA PPPASTAFVV SGLAGAAVGS
 HCV1480    SFTAAVTSPL TTHQTLLFNI LGGWVASQIA PPTAATAFVV SGLAGAAVGS
 HCVPOLYP   AFTASITSPL TTQHTLLFNI LGGWVAAQLA PPSAASAFVG SGMAGAAVGN
   HD_1     AFTASITSPL TTQSTLLFDI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
 HPCCGAA    AFTAAVTSPL TTGQTLLFNI LGGWVAAQLA APGAATAFVG AGIARAAVGS
   HPCFG    AFTASVTSPL TTNQTMFFNI LGGWVATHLA GPAASSAFVV SGLAGAALDS
HPCGENANTI  AFTASITSPL TTQSTLLFNI LGGWVAAQLA PPGAASAFVG AGIAGAAVGG
 HPCGENOM   AFTASITSPL TTQHTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
 HPCHUMR    AFTASITSPL TTQSTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
   HPCJ     AFTSSITSPL TTQSTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
  HPCJCG    AFTASITSPL TTQNTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
 HPCJK046   SFSASLTSPL TTAQTLLLNV LGGWVASQLA TPVPATAFVV SGLAGAAIGS
 HPCJK049   AFTASVTSPL TTNQTLLFNI MGGWVASNLA PPPASTAFVV SGLAGAAIGS
  HPCJTA    AFTASITSPL TTQNTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
  HPCJTB    AFTASITSPL TTQNTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAIGS
  HPCK3A    AFTASVTSPL TTNQTMFFNI LGGWVATHLA GPQASSAFVV SGLAGAAIGS
 HPCPLYPRE  AFTAAVTSPL TTSQTLLFNI LGGWVAAQLA APGAATAFVG AGLAGAAIGG
  HPCPOLP   AFSAALTSPL STSTTILLNI LGGWLASQIA PPAGATGFVV SGLVGAAVGS
   HPCPP    AFTASITSPL TTQYTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
 HPCUNKCD   AFTASITSPL TTQHTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGS
   MKC1A    AFTASITSPL TTQYTLLFNI LGGWVAAQLA PPSAASAFVG AGIAGAAVGT
   NDM59    AFSAALTSPL STSTTILLNI LGGWLASQIA PPAGATGFVV SGLVGAAVGS
    NZLI    AFTASVTSPL TTNQTMFFNI LGGWVATHLA GPQSSSAFVV SGLAGAAIGG
    SA13    SFTAAVTSPL TTQQTLLFNI LGGWVASQIA PPTAATAFVV SGMAGAAVGS
    Th580   SFTASLTSPL RTSQTLLLNI LGGWIASQVA PPSASTAFVV SGLAGATVAS
Type_3a_CB  AFTASVTSPL TTNQTMFFNI LGGWVATHLA GPQSSSAFVV SGLAGAAIGG
  TypeV_D   AFTASVTSPL TTNQTMFFNI LGGWVATHLA GPQSSSAFVV SGLAGAAIGG
   VN004    SFSASLTSPL STHQTLLLNI LGGWVASQLA NPTASTAFVV SGLAGAAVGS
   VN235    SFSAALTSPL STSTTLLLNI LGGWVASQLA PPTASTAFVV SGLAGAAVGS
   VN405    SFSASLTSPL STSTTLLLNI LGGWVASQLA NPTASTAFVV SGLAGATVGS

            1851                                                1900
   BEBE1    IGLGKILVDV LAGYGAGISG ALVAFKIMSG EKPSVEDVVN LLPAILSPGA
   D89815   IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG DMPSTEDLVN LLPAILSPGA
ED43type_4  VGLGKILVDI LPGYGAGVRG AVVTFKIMSG EMPSTEDLVN LLPAILSPGA
   HC_C2    IGLGKVLVDI LAGYGAGVAG ALVAFKIMSG ETPSAEDLVN LLPAILSPGA
   HC_G9    VGLGKVLVDI LAGYGAGVAG ALVAFKIMSG ETPSAEDLVN LLPAILSPGA
  HCU16326  IGLGKVLVDI LAGYGAGVAG ALVAFKIMSG EAPTAEDLVN LLPAILSPGA
 HCV_H_CMR  VGLGKVLVDI LAGYGAGVAG ALVAFKIMSG EMPSAEDMVN LLPAILSPGA
   HCV_J1   VGLGRVLVDI LAGYGAGVAG ALVAFKIMSG EVPSTEDLVN LLPAILSPGA
  HCV_J483  IGLGKVLVDI LAGYGAGVAG ALVAFKIMSG ELPSTEDLVN LLPAILSPGA
   HCV_J8   IGLGKILVDV LAGYGAGISG ALVAFKVMSG EVPSTEDLVN LLPAILSPGA
  HCV_JK1   IGLGKVLVDI LAGYGAGVAG ALVAFKIMSG EKPTVEDVVN LLPAILSPGA
   HCV_JS   IGLGKVLVDI LAGYGAGVAG ALVAFKGMSG EMPSTEDLVN LLPAILSPGA
 HCV_K1_R1  IGLGKVLVDI LAGYGAGVAG ALVAFKIMSG EMPATEDLVN LLPAILSPGA
 HCV_K1_R2  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EMPSTEDLVN LLPAILSPGA
 HCV_K1_R3  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG DMPSTEDLVN LLPAILSPGA
 HCV_K1_S1  IGLGKVLVDI LAGYGAGVAG ALVDFKVMSG EMPSAEDIVN LLPAILSPGA
 HCV_K1_S2  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EMPSTEDLVN LLPAILSPGA
 HCV_K1_S3  IGLGKVLVDI LAGYGAGVAG ALVDFKVMSG DMPSTEDLVN LLPAILSPGA
   HCV_L2   IGLGKVLVDI LAGYGAGVAG ALVDFKVMSG EMPSAEDIVN LLPAILSPGA
   HCV_N    IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EMPSTEDLVN LLPAILSPGA
 HCV12083   IRLGRVLVDV LAGYGAGVSG ALVAFKIMSG EAPSAEDLIN LLPAILSPGA
 HCV1480    IGLGRVLIDI LAGYGTGVAG ALVAFKIMCG ECPSTEDMVN LLPALLSPGV
 HCVPOLYP   IGLGKVLVDI LAGYGAGVAG ALVAFKIMCG ERPTAEELVN LLPSILCPGA
   HD_1     IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EMPSTEDLVN LLPAILSPGA
 HPCCGAA    IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EVPSTEDLIN LLPAILSPGA
 HPCCGAA    VGLGKVLVDI LAGYGAGVAG ALVAFKIMSG EVPSTEDLVN LLPAILSPGA
```

```
          HPCFG  IGIGRVLLDV LAGYGAGVSG ALVAFKIMGG ELPTTEDMVN LLPAILSPGA
     HPCGENANTI  IGLGKVLVDM VAGYGAGVAG ALVAFKVMSG EMPSTEDLVN LLPAILSPGA
        HPCGENOM  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EMPSTEDLVN LLPAILSPGA
        HPCHUMR  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EMPSTEDLVN LLPAILSPGA
           HPCJ  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EVPSTEDLVN LLPAILSPGA
          HPCJCG  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EMPSTEDLVN LLPAILSPGA
        HPCJK046  IGLGKVIVDI LAGYGAGVSG ALVAFKIMSG ETPSVEDMVN LLPAILSPGA
        HPCJK049  IGLGKVLLDI LAGYGAGVAG ALVAFKIMGG EMPSTEDMVN LLPALLSPGA
          HPCJTA  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EAPSAEDLVN LLPAILSPGA
          HPCJTB  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG EAPSAEDLVN LLPAILSPGA
          HPCK3A  IGLGRVLLDI LAGYGAGVSG ALVAFKIMGG EPPTTEDMVN LLPAILSPGA
       HPCPLYPRE  VGLGKVLIDI LAGYGAGVAG ALVAFKIMSG EVPSTEDLVN LLPAILSPGA
          HPCPOLP  IGLGKVLVDI LAGYGAGISG ALVAFKIMSG EKPSMEDVVN LLPGILSPGA
           HPCPP  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG DMPSTEDLVN LLPAILSPGA
        HPCUNKCD  IGLGKVLVDI LAGYGAGVAG ALVAFKIMSG EMPSAEDMVN LLPAILSPGA
           MKC1A  IGLGKVLVDI LAGYGAGVAG ALVAFKVMSG DMPSTEDLVN LLPAILSPGA
           NDM59  IGLGKVLVDI LAGYGAGISG ALVAFKIMSG EKPSMEDVIN LLPGILSPGA
            NZLI  IGLGRVLLDI LAGYGAGVSG ALVAFKIMGG ECPTAEDMVN LLPAILSPGA
            SA13  IGLGRVLIDI LAGYGAGVAG ALVAFKIMCG EKPTAEDLVN LLPSILCPGA
            Th580  IGLGRVIVDI LAGYGAGVAG ALVAFKIMSG ECPSTEDMVN LLPALLSPGA
       Type_3a_CB  IGLGRVLLDI LAGYGAGVSG ALVAFKIMGG ELPTAEDMVN LLPAILSPGA
          TypeV_D  IGLGRVLLDI LAGYGAGVSG ALVAFKIMGG ELPTTEDLVN LLPAILSPGA
            VN004  IGLGRVIVDV LAGYGAGVSG ALVAFKIMCG ETPSAEDMVN LLPALLSPGA
            VN235  IGLGKVIIDI LAGYGAGVSG ALVAFKIMSG EAPAVEDMVN LLPALLSPGA
            VN405  IGLGRVLVDI IAGYGAGVSG ALVAFKIMSG ETPSAEDMVN LLPALLSPGA

                   1901                                                1950
            BEBE1  LVVGVICAAI LRRHVGQGEG AVQWMNRLIA FASRGNHVAP THYVAESDAS
            D89815  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
       ED43type_4  LVVEVVCPAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
             HC_C2  LVVGVVCAAI QRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
             HC_G9  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
          HCU16326  LVVGIVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAS
          HCV_H_CMR  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP RHYVPESEPA
             HCV_J1  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
           HCV_J483  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
             HCV_J8  LVVGVICAAI LRRHVGQGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
            HCV_JK1  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVAP THYVVESDAS
             HCV_JS  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
          HCV_K1_R1  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
          HCV_K1_R2  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
          HCV_K1_R3  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
          HCV_K1_S1  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
          HCV_K1_S2  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
          HCV_K1_S3  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
             HCV_L2  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
              HCV_N  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHDSP THYVPESDAA
          HCV12083  ALVGVVCAAI LRRHVGPAEG ANQWMNRLIA FASRGNHVSP THYVPESDAA
           HCV1480  LVVGVICAAV LRRHIGPGEG AVQWMNRLIA FASRGNHVSP THYVPETDAS
          HCVPOLYP  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHGSP THYVPETDAS
              HD_1  LVVGVVCAAI LRGHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
          HPCCGAA  LAVGVVFASI LRRRVGPGEG AVQWMNRLIA FAFAGNHVSP THYVPESDAA
            HPCFG  LVVGVICAAV LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
     HPCGENANTI  LVVGVVCAAI LRRHVDPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
        HPCGENOM  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
        HPCHUMR  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
           HPCJ  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
          HPCJCG  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
        HPCJK046  LVVGVVRAAI LRRHVGPSEG AAQWMNRLIA FASRGNHVSP THYVPETDAS
        HPCJK049  LVVGVICAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVAP THYVPESDAA
          HPCJTA  LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
```

```
HPCJTB   LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
HPCK3A   LVVGVICAAI LRRHVGPGEG PVQWMNRLIA FASRGNHVSP AHYVPESDAA
HPCPLYPRE LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
HPCPOLP  LVVGVICAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVAP THYVTESDAS
HPCPP    LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
HPCUNKCD LVVGIVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP RHYVPESEPA
MKC1A    LVVGVVCAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
NDM59    LVVGVICAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVAP THYVTESDAS
NZLI     LVVGVICAAV LRRHIGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
SA13     LVVGVVCAAI LRRHVGPSEG ANQWMNRLIA FASRGNHVSP THYVPETDAS
Th580    LVVGVICAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPETDAS
Type_3a_CB LVVGVICAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
TypeV_D  LVVGVICAAI LRRHVGPGEG AVQWMNRLIA FASRGNHVSP THYVPESDAA
VN004    LVVGVVCAAI LRRHAGPSEG ATQWMNRLIA FASRGNHVSP THYVPETDTS
VN235    LVVGVVCAAV LRRHVGPSEG ATQWMNRLIA FASRGNHVSP THYVPETDAS
VN405    LVVGVVCAAI LRRHAGPAEG ATQWMNRLIA FASRGNHVSP THYVPETDTS

         1951                                                2000
BEBE1    QRVTQLLGSL TITSLLRRLH QWITEDCPVP CSGSWLRDVW DWVCSILIDF
D89815   ARVTQILSNL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLADF
ED43type_4 RRVTTILSSL TVTSLLRRLH KWINEDCSTP CAESWLWEVW DWVLHVLSDF
HC_C2    ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDIW DWICSVLTDF
HC_G9    VRVTHILTSL TVTQLLKRLH VWISSDCTAP CAGSWLKDVW DWICEVLSDF
HCU16326 ARVTQILSSL TITQLLKRLH QWINEDCSTP CSSSWLREIW DWICTVLTDF
HCV_H_CMR ARVTAILSSL TVTQLLRRLH QWISSECTTP CSGSWLRDIW DWICEVLSDF
HCV_J1   ARVTAILSSL TVTQLLRRLH QWLSSESTTP CSGSWLRDIW DWICEVLSDF
HCV_J483 ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HCV_J8   QRVTQVLSSL TITSLLRRLH AWITEDCPVP CSGSWLQDIW DWVCSILTDF
HCV_JK1  ARVTKILSSL TITQRLRRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HCV_JS   ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDIW DWICTVLTDF
HCV_K1_R1 ARVTQILSSL TITQLLRRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HCV_K1_R2 ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLSDF
HCV_K1_R3 VRVTQILSNL TITQLLKRLH QWISEDCSTP CSGSWLRDVW DWICTVLTDF
HCV_K1_S1 ARVTQILSSL TITQLLRRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HCV_K1_S2 ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HCV_K1_S3 VRVTQILSSL TITQLLKRLH QWISEDCSTP CSGSWLRDVW DWICTVLTDF
HCV_L2   ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HCV_N    ARVTQVLSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWVCTVLSDF
HCV12083 KNVTQILTSL TITSLLRRLH QWVNEDTATP CATSWLRDVW DWVCTVLSDF
HCV1480  AKVTQLLSSL TVTSLLKRLH TWIGEDYSTP CDGTWLRAIW DWVCTALTDF
HCVPOLYP ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HD_1     ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HPCCGAA  ARVTAILSSL TVTQLLRRLH QWISSECTTP CSGSWLRDIW DWICEVLSDF
HPCFG    AKVTALLSSL TVTRLLRRLH QWINEDYPSP CNGDWLHDIW DWVCIVLSDF
HPCGENANTI ARVTQILSGL TITQLLRRLH QWINEDCSTP CSGSWLRDVW DWICTVLADF
HPCGENOM ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HPCHUMR  ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HPCJ     QRVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HPCJCG   ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLKDVW DWICTVLSDF
HPCJK046 RAVTNILSSL TITSLLRKLH HWITEDYATP CGSTWLRDIW DWVCTVLSDF
HPCJK049 AKVTALLSSL TVTQLLRRLH QWINEDYPTP CDGNWLYDIW NWVCTVLADF
HPCJTA   ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLKDVW DWICTVLTDF
HPCJTB   ARVTQILSSL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLTDF
HPCK3A   ARVTALLSSL TVTSLLRRLH QWINEDYPSP CSGDWLRIIW DWVCSVVSDF
HPCPLYPRE ARVTAILSSL TVTQLLRRLH QWISSECTTP CSGSWLRDIW DWICEVLSDF
HPCPOLP  QRVTQLLGSL TITSLLRRLH NWITEDCPIP CSGSWLRDVW DWVCTILTDF
HPCPP    ARVTQILSNL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLADF
HPCUNKCD ARVTQILSSL TITQLLKRLH QWINEDCSTP CSSSWLREIW DWICTVLTDF
MKC1A    ARVTQILSNL TITQLLKRLH QWINEDCSTP CSGSWLRDVW DWICTVLADF
NDM59    QRVTQLLGSL TITSLLRRLH NWITEDCPIP CAGSWLREVW DWVCTILTDF
NZLI     ARVTALLSSL TVTSLLRRLH QWINEDYPSP CSDDWLRTIW DWVCSVLADF
```

```
          SA13  AKVTQLLSSL TVTSLLKRLH TWIGEDYSTP CDGTWLRAIW DWVCTALTDF
          Th580 NKVTQILSSL TITSLLRRLH QWIHEDTSTP CASSWLRDVW DWVCTVLSDF
      Type_3a_CB ARVTALLSSL TVTSLLRRLH QWINEDYPSP CSDDWLRIIW DWVCSVLSDF
        TypeV_D AKVTALLSSL TVTSLLRRLH QWINEDYPSP CSDDWLRIIW DWVCSVLADF
          VN004 RQIMTILSSL TVTSLLRKLH EWINTDWSTP CSSSWLRDIW DWVCEVLSDF
          VN235 RAVTTILSSL TITSLLRRLH EWISGDWSAP CSCSWLRDIW DWVCTVLSDF
          VN405 RQVMAILSSL TVTSLLRKLH EWINSDWSTP CSGSWLRDIW DWVCTVLSDF

                2001                                                2050
          BEBE1 KNWLSAKLFP RLPGIPFISC QKGYRGTWAG TGIMTTRCPC GANITGNVRL
          D89815 KTWLQSKLLP RLPGVPFFSC QRGYKGVWRG DGIMYTTCPC GAQITGHVKN
      ED43type_4 KTCLKAKFVP LMPGIPLLSW PRGYKGEWRG DGVMHTTCPC GADLAGHIKN
          HC_C2 KTWLQSKLLP RLPGVPFFSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
          HC_G9 KSWLKAKLMP QLPGIPFVSC QRGYRGVWRG DGIMQTTCPC GAQITGHVKN
        HCU16326 KTWLQSKLLP RLPGVPFFSC QRGYKGVWRG EGIMHARCPC GADITGHVKN
        HCV_H_CMR KTWLKAKLMP QLPGIPFVSC QRGYRGVWRG DGIMHTTCPC GAQITGHVKN
          HCV_J1 KTWLKTKLMP HLPGIPFVSC QHGYKGVWRG DGIMHTRCHC GAEITGHVKN
        HCV_J483 KTWLQSKLLP RLPGVPFLSC QRGYKGVWRG DGIMHTRCHC GAEITGHVKN
          HCV_J8 KNWLSSKLLP KMPGIPFISC QKGYKGVWAG TGVMTTRCPC GANISGHVRM
        HCV_JK1 KTWLQSKLLP RLPGDPFFSC QRGYRGVWRG DGVMQTTCPC GAQITGHVKN
          HCV_JS KTWLKSKLMP RLPGVPFFSC QRGYRGVWRG DGIMHTTCPC GAQITGHVKN
        HCV_K1_R1 KTWLQSKLLP RLPGVPFFSC QRGYRGVWRG DGIMQTTCPC GAQITGHVKN
        HCV_K1_R2 KTWLQSRVLP RLPGVPFLSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
        HCV_K1_R3 KTWLQSKLLP RLPGIPFFSC QRGYKGVWRG DGVMHTICPC GAQIAGHFKN
        HCV_K1_S1 KTWLQSKLLP RLPGVPFFSC QRGYRGVWRG DGIMQTTCPC GAQITGHVKN
        HCV_K1_S2 KTWLQSRVLP RLPGVPFLSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
        HCV_K1_S3 KTWLQSKLLP RLPGIPFFSC QRGYKGVWRG DGVMHTICPC GAQIAGHFKN
          HCV_L2 KTWLQSKLLP RLPGVPFFSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
          HCV_N KTWLQSKLLP RLPGVPFLSC QRGYKGVWRG DGIMHTTCPC GAQIAGHVKN
        HCV12083 KVWLQAKLFP RLPGIPFLSC QAGYRGVWAG DGVCHTTCTC GAVIAGHVKN
        HCV1480 KAWLQAKLLP QLPGVPFFSC QKGYKGVWRG DGVNSTKCPC GATISGHVKN
        HCVPOLYP KTWLQSKLLP RLPGIPFYSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
          HD_1 KTWLQSKLLP RLPGVPFLSC QRGYRGVWRG DGIMHTTCPC GAQMAGHVKN
        HPCCGAA KTWLKAKLMP QLPGIPFVSC QRGYRGVWRG DGIMHTRCHC GAEITGHVKN
          HPCFG KTWLSAKIMP KVPGIPFLSC QKGYKGVWRG DGVMTTRCPC GEDFTGHVRN
      HPCGENANTI KTWLQSKLLP RLPGVPFFSC QRGYKGVWRG DGIMQTTCPC GAQLTGHVKN
        HPCGENOM KTWLQSKLLP RLPGVPFLSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
        HPCHUMR KTWLQSKLLP QLPGVPFFSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
          HPCJ KTWLQSKLLP RLPGVPFFSC QRGYKGVWRG EGIMQTTCPC GAQIAGHVKN
        HPCJCG KTWLQSKLLP RLPGLPFLSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
        HPCJK046 RVWLKSKLMP SLPGVPFFSC QRGYRGTWRG DGICNTTCPC GASIAGHVKN
        HPCJK049 KLWLGAKILP KMPGIPFLSC QKGYRGTWRG DGVVSTRCPC GALLSGHVKN
          HPCJTA KTWLQSKLLP KLPGVPFFSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
          HPCJTB KTWLQSKLLP QLPGVPFFSC QRGYKGVWRG DGIMQTTCPC GAQITGHVKN
          HPCK3A KTWLSAKIMP ALPGLPFISC QKGYKGVWRG DGVMSTRCPC GASIAGHVKN
      HPCPLYPRE KTWLKAKLMP QLPGIPFVSC QRGYKGVWRV DGIMHTRCHC GAEITGHVKN
        HPCPOLP KNWLTSKLFP KMPGLPFISC QKGYKGVWAG TGIMTTRCPC GANISGNVRL
          HPCPP KTWLQSKLLP RLPGVPFFSC QRGYKGVWRG DGIMYTTCPC GAQITGHVKN
        HPCUNKCD KTWLQSKLLP RLPGVPFFSC QRGYKGVWRG DGIMHTTCPC GAQITGHVKN
          MKC1A KTWLQSKLLP RLPGVPFFSC QRGYKGVWRG DGIMYTTCPC GAQITGHVKN
          NDM59 KNWLTSKLFP KMPGLPFISC QKGYRGVWAG TGIMTTRCPC GANISGNVRL
          NZLI KAWLSAKIMP ALPGLPFISC QKGYKGVWRG DGVMSTRCPC GAAITGHVKN
          SA13 KAWLQAKLLP QLPGVPFLSC QRGYRGVWRG DGVNSTKCPC GATISGHVKN
          Th580 KTWLKAKITP RIPGIPFISC QAGYRGVWAG DGVCHTTCSC GAQIAGHVKN
      Type_3a_CB KSWLSAKIMP ALPGLPFISC QKGYKGVWRG DGVMSTRCPC GATITGHVKN
        TypeV_D KTWLSAKIMP ALPGLPFISC QKGYKGVWRG DGVTTTRCPC GATITGHVKN
          VN004 KTWLKAKLVP ALPGVPFLSC QRGFRGTWRG DGICHTTCPC GSEITGHVKN
          VN235 KTWLRAKLVP TLPGIPFISC QRGFRGVWRG DGVNYTTCSC GANITGHVKN
          VN405 KVWLKSKLVP ALPGVPFLSC QRGFRGVWRG DGICRTTCPC GADIVGHVKN

                2051                                                2100
```

```
          BEBE1  GTMRISGPKT CLNTWQGTFP INCYTEGSCV PKPAPNFKTA IWRVAASEYA
          D89815 GSMRIVGPRT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
       ED43type_4 GSMRITGPKT CSNTWHGTFP INAYTTGPGV PIPAPNYKFA LWRVSAEDYV
          HC_C2  GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
          HC_G9  GSMRIVGPKT CSNTWRGSFP INAHTTGPCT PSPAPNYTFA LWRVSAEEYV
        HCU16326 GSMRIVGPKT CSNTWYGTFP INAYTTGPCT PSPAPNYSKA LWRVAAEEYV
        HCV_H_CMR GTMRIVGPRT CRNMWSGTFP INAYTTGPCT PLPAPNYKFA LWRVSAEEYV
          HCV_J1 GTMRIVGPKT CRNMWSGTFP INAYTTGPCT PLPAPNYTFA LWRVSAEEYV
        HCV_J483 GSMRIVGPRT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
          HCV_J8 GTMKITGPKT CLNLWQGTFP INCYTEGPCV PKPPPNYKTA IWRVAASEYV
        HCV_JK1 GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
          HCV_JS GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
        HCV_K1_R1 GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
        HCV_K1_R2 GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
        HCV_K1_R3 GSMKIVGPKT CSNTWDGTFP INGYTTGSST PTPASNYSKA LWRVVFEEYV
        HCV_K1_S1 GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
        HCV_K1_S2 GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
        HCV_K1_S3 GSMKIVGPKT CSNTWDGTFP INGYTTGSST PTPASNYSKA LWRVVFEEYV
          HCV_L2 GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PAPTPNYSRA LWRVAAEEYV
          HCV_N  GSMRIIGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSKA LWRVAAEEYV
        HCV12083 GTMKITGPKT CSNTWHGTFP INATTTGPST PRPAPNYQRA LWRVSAEDYV
          HCV1480 GTMRIVGPKL CSNTWQGTFP INATTTGPSV PAPAPNYKFA LWRVGAADYA
        HCVPOLYP GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
          HD_1   GSMRIVGPKT CSNTWYGSFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
        HPCCGAA  GTMRIVGPRT CKNMWSGTFF INAYTTGPCT PLPAPNYKFA LWRVSAEEYV
          HPCFG  GSMRIAGSGL CANMWHGTFP INEYTTGPST PVPAHNYSRA LWRVTSDSYV
      HPCGENANTI GSMRIWGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
        HPCGENOM GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAPEEYV
        HPCHUMR GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
          HPCJ   GSMRIVGPRT CSNTWHGTFP INAYTTGPCS PSPAPNYSRA LWRVAAEEYV
        HPCJCG   GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
        HPCJK046 GTMRIVGPRT CSNVWNGTFP INATTTGPSI PIPAPNYKKA LWRVSATEYV
        HPCJK049 GTMRLVGPRW CANTWHGTFP INGYTTGPST PAPSYAYSRA LWRVASDSYV
          HPCJTA GSMRIVGPKT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
          HPCJTB GSMRIVGPKT CSNMWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
          HPCK3A GSMRLAGPRT CANMCHGTFP INEYTTGPST PCPPPNYTRA LWRVAANSYV
        HPCPLYPRE GTMRIVGPRT CRNMWSGTFP INAYTTGPCT PLPAPNYTFA LWRVSAEEYV
          HPCPOLP GSMRITGPKT CMNIWQGTFP INCYTEGQCV PKPAPNFKIA IWRVAASEYA
          HPCPP  GSMRIVGPRT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
        HPCUNKCD GSMRIVGPKT CSNTWYGTFP INAYTTGPCT PSPAPNYSKA LWRVAAEEYV
          MKC1A  GSMRIVGPRT CSNTWHGTFP INAYTTGPCT PSPAPNYSRA LWRVAAEEYV
          NDM59  GSMRITGPKT CMNTWQGTFP INCYTEGQCV PKPAPNFKTA IWRVAASEYA
          NZLI   GSMRLAGPRT CANMWHGTFP INEYTTGPST PCPSPNYTRA LWRVAANSYV
          SA13   GTMRIVGPKL CSNTWHGTFP INATTTGPSV PAPAPNYKFA LWRVGAADYA
          Th580  GSMKITGPRM CSNTWHGTFP INATTTSPSV PVPAPNYKRA LWRVSAEEYV
      Type_3a_CB GSMRLAGPRT CANMWHGTFP INEYTTGPST PCPSPNYTRA LWRVAANSYV
        TypeV_D  GSMRLAGPRT CANMWYGTFP INEYTTGPST PCPSPNYTRA LWRVAANSYV
          VN004  GTMKISGPRW CSNVSHRTFP INATTTGPSV PIPEPNYTRA LWRVSAEEYV
          VN235  GSMKIVGPKM CSNVWNNRFP INAITTGPSV PVPEPNYHKA LWRVSAEDYV
          VN405  GSMRISGSRW CSNIWHGTFP INATTTGPSV PIPEPNYKRA LWRVSAEEYV

                 2101                                                2150
          BEBE1  EVTQHDSHAY VTGLTADNLK VPCQLPCPEF FSWVDGVQIH RFAPTPKAFM
          D89815 EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTELDGVRLH RYAPACKPLL
       ED43type_4 EVRRVGDFHY VTGVTQDNIK FPCQVPAPEL FTEVDGIRIH RHAPKCKPLL
          HC_C2  EVTRVGDFHY ITGMTTDNIK CPCQVPAPEF FTEVDGVRLH RYAPACKPVL
          HC_G9  EVRRLGDFHY ITGVTTDKIK CPCQVPSPEF FTEVDGVRLH RYAPPCKPLL
        HCU16326 EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVRLH RYAPACRPLL
        HCV_H_CMR EIRRVGDFHY VSGMTTDNLK CPCQIPSPEF FTELDGVRLH RFAPPCKPLL
          HCV_J1 EIRRVGDFHY VTGMTTDNLK CPCQVPSPEF FTELDGVRLH RFAPPCKPLL
        HCV_J483 EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVRLH RYAPACKPLL
```

```
    HCV_J8  EVTQHGSFSY VTGLTSDNLK VPCQVPAPEF FSWVDGVQIH RFAPVPGPFF
   HCV_JK1  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVRLH RYAPACKPLL
    HCV_JS  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVQLH RYAPACKPLL
 HCV_K1_R1  EVTRVGDFHY VTGMTTDNVK CPCQVPPPEF FTEVDGVRLH RNAPACGPLL
 HCV_K1_R2  EVTRVGDFHY VTGMTTDNLK CPCQVPAPEF FKELDGVRLH RYAPASKPLL
 HCV_K1_R3  EVTRVGDFHY VTGMTTDNVK CPCQVPPPEF FTELDGVRLH RYAPVSKPLL
 HCV_K1_S1  EVTRVGDFHY VTGMTTDNVK CPCQVPPPEF FTEVDGVRLH RNAPACGPLL
 HCV_K1_S2  EVTRVGDFHY VTGMTTDNLK CPCQVPAPEF FKELDGVRLH RYAPASKPLL
 HCV_K1_S3  EVTRVGDFHY VTGMTTDNVK CPCQVPPPEF FTELDGVRLH RYAPVSKPLL
    HCV_L2  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVRLH RYAPACKTLL
     HCV_N  EVTRVGDFHY VTGITTDNVK CPCQVPAPEF FTEVDGVRLH RYAPVCKPLL
  HCV12083  EVRRLGDCHY VVGVTAEGLK CPCQVPAPEF FTEVDGVRIH RYAPPCKPLL
   HCV1480  EVRRVGDYHY ITGVTQDNLK CPCQVPSPEF FTELDGVRIH RFAPPCNPLL
  HCVPOLYP  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVRLH RYAPACKPLL
      HD_1  EVTRVGDFHY VTGMTTDNIK CPCQVPAPEF FTEVDGVRLH RYAPACKPLL
   HPCCGAA  EIRRVGDFHY VSGMTTDNLK CPCQIPSPEF FTELDGVRLH RFAPPCKPLL
     HPCFG  EVRRVGDTHY VVGATNDGLK IPCQVPAPEF FTELDGVRLH RYAPPCKPLL
HPCGENANTI  EVRRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVRLH RYAPACKPLL
  HPCGENOM  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVRLH RYAPACKPLL
   HPCHUMR  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FSEVDGVRLH RYAPACRPLL
      HPCJ  EVTRVGDFHY VTGVTTDNVK CPCQVPAPEF FTELDGVRLH RYAPACKPLL
    HPCJCG  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVRLH RYAPVCKPLL
  HPCJK046  EVVRVGDSHY ITGVTAENTK CPCQVPAPEF FTEVDGVRLH RYAPECKPIL
  HPCJK049  EVRKVGDFHY VTGTTDDGLK CPCQVPLPEF FTELDGVRLH RYAPVCRPLL
    HPCJTA  EITRVGDFHY VTGMTTDNVK CPCQVPAPEF FTELDGVRLH RYAPACRPLL
    HPCJTB  EITRVGDFHY VTGXTTDNVK CPCQVPAPEF FTELDGVRLH RYAPACRPLL
    HPCK3A  EVRRVGDFHY ITGATEDGLK CPCQVPATEF FTEVDGVRIH RYAPPCRPLL
 HPCPLYPRE  EIRQVGDFHY VTGMTTDNLK CPCQVPSPEF FTELDGVRLH RFAPPCKPLL
   HPCPOLP  EVTQHGSYHY ITGLTTDNLK VPCQLPSPEF FSWVDGVQIH RFAPIPKPFF
     HPCPP  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTELDGVRLH RYAPACKPLL
  HPCUNKCD  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTEVDGVRLH RYAPACRPLL
     MKC1A  EVTRVGDFHY VTGMTTDNVK CPCQVPAPEF FTELDGVRLH RYAPACKPLL
     NDM59  EVTQHGSYSY ITGLTTDSLK VPCQLPSPEF FSWVDGVQIH RFAPTPKPFF
      NZLI  EVRRVGDFHY ITGATEDELK CPCQVPAAEF FTEVDGVRLH RYAPPCKPLL
      SA13  EVRRVGDYHY ITGVTQDNLK CPCQVPSPEF FTELDGVRIH RYAPPCNPLL
     Th580  EVERHGDRHY VVGVTADGLK CPCQVPGPEF FTEVDGVRIH RYAPPCKPLL
 Type_3a_CB  EVRRVGDFHY ITGATEDELK CPCQVPAAEF FTEVDGVRLH RYAPPCKPLL
   TypeV_D  EVRRVGDFHY ITGATEDELK CPCQVPAAEF FTEVDGVRLH RYAPPCKPLL
     VN004  EVKRVGDSHF VVGATTDNLK CPCQVPAPEF FTEVDGVRLH RYAPRCKPLL
     VN235  EVVRVNDHHY IVGATADNLK CPCQVPAPEF FTEVDGVRLH RFAPPCRPLM
     VN405  EVARVGDSHF VVGATNQDLK CPCQVPAPEF FTEVDGVRLH RFAPACKPLL

            2151                                              2200
     BEBE1  RDEVSFSVGL NSYVVGSQLP CEPEPDTEVL ASMLTDPSHI TAEAAARRLA
     D89815  RDEVTFQVGL NQYTVGSQLP CEPEPDVTVV TSMLTDPSHI TAEAARRRLA
ED43type_4  RDEVSFSVGL NSFVVGSQLP CEPEPDVAVL TSMLTDPSHI TAESARRRLA
      HC_C2  REEVDFQVGL NQYPVGSQLP CEPEPDVAVL TSMLTDPSHI TAEAAKRRLA
      HC_G9  RDEVTFSIGL NEYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAARRLN
   HCU16326  REEVVFQVGL HQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
  HCV_H_CMR  REEVSFRVGL HEYPVGSQLP CEPEPDVAVL TSMLTDPSHI TAEAAGRRLA
     HCV_J1  REEVSFRVGL HDYPVGSQLP CEPEPDVAVL TSMLTDPSHI TAAAAGRRLA
   HCV_J483  REDVAFQVGL NQYLVGSQLP CEPEPDVTVL TSMLTDPSHI TAETAKRRLA
     HCV_J8  RDEVTFTVGL NSFVVGSQLP CDPEPDTEVL ASMLTDPSHI TAEAAARRLA
    HCV_JK1  RDEVTFQVGL NQFPVGSQLP CEPEPDVTVL TSMLTDPSHI TAETAKRRLA
     HCV_JS  RDEVTFQVGL NQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAEAAKRRLA
  HCV_K1_R1  REEVTFQVGL NQYLVGSQLP CEPEPDVTVL TSMLTDPSHI TGEAAKRRLA
  HCV_K1_R2  RDEVTFQVGL NQYVVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
  HCV_K1_R3  RDEVTFQVGL NRYPVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
  HCV_K1_S1  REEVTFQVGL NQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TGETAKRRLA
  HCV_K1_S2  RDEVTFQVGL NQYVVGSQLP CEPEPDVAVL TSMLTDPSHI TAEAAKRRLA
  HCV_K1_S3  RDEVTFQVGL NRYAVGSQLP CEPEPDVTVI TSMLTDPSHI TAETAKRRLA
```

```
HCV_L2        REEVTFQVGL NQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
HCV_N         RDEVVFQVGL NQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAEAAKRRLA
HCV12083      RDEVTFSVGL SNYAVGSQLP CEPEPDVTVV TSMLTDPTHI TAETAARRLK
HCV1480       REEVTFSVGL HSYVVGSQLP CEPEPDVTVL TSMLSDPAHI TAETAKRRLN
HCVPOLYP      RDEVTFQVGL NQYVVGSQLP CEPEPDVVVV TSMLTDPSHI TAETAKRRLD
HD_1          RDEVSFQVGL NHYPVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
HPCCGAA       REEVSFRVGL HEYPVGSQLP CEPEPDVAVL TSMLTDPSHI TAEAAGRRLA
HPCFG         RDEITFSVGL HSYANGSQLS CEPEPDVAVL TSMLRDPAHI TAATAARRLA
HPCGENANTI    REEVSFQVGL NQYVVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
HPCGENOM      REEVVFQVGL NQYLVGSQLP CEPEPDVTVL TSMLTDPSHI TAETAKRRLA
HPCHUMR       REEVTFQVGL NQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
HPCJ          RDEVSFQVGL NQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
HPCJCG        REEVVFQVGL NQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
HPCJK046      RDEVTFTVGL STYVVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
HPCJK049      RDDVTFTVGL NSYVIGSQLP CEPEPDVLVV TSMLRDPDHI TAEEASRRLK
HPCJTA        REDVTFQVGL NQYLVGSQLP CEPEPDVAVV TSMLQDPSHI TVETAKRRLD
HPCJTB        REDVTFQVGL NQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
HPCK3A        RDEITFMVGL NSYAIGSQLP CEPEPDVSVL TSMLRDPSHI TAETAARRLA
HPCPLYPRE     REEVSFRVGL HEYPVGSQLP CEPEPDVAVL TSMLTDPSHI TAEAAGRRLA
HPCPOLP       RDEVSFCVGL NSFVVGSQLP CDPEPDTDVL TSMLTDPSHI TAETAARRLA
HPCPP         RDEVTFQVGL NQYTVGSQLP CEPEPDVTVV TSMLTDPSHI TAEAARRRLA
HPCUNKCD      REEVVFQVGL HQYLVGSQLP CEPEPDVAVL TSMLTDPSHI TAETAKRRLA
MKC1A         RDEVTFQVGL NQYTVGSQLP CEPEPDVTVV TSMLTDPSHI TAEAARRRLA
NDM59         RDEVSFCVGL NSFVVGSQLP CDPEPDADVL TSMLTDPSHI TAEAAARRLA
NZLI          RDDITFMVGL HSYTIGSQLP CEPEPDVSVL TSMLRDPSHI TAETAARRLA
SA13          REEVCFSVGL HSFVVGSQLP CEPEPDVTVL TSMLSDPAHI TAETAKRRLD
Th580         RDEVSFSVGL LEFVVGSQLP CEPEPDVTVV TSMLTDPSHI TAETASRRLK
Type_3a_CB    REEITFSVGL NSYTIGSQLP CEPEPDVSVL TSMLRDPSHI TAETAARRLA
TypeV_D       RDDITFMVGL NSYAIGSQLP CEPEPDVSVL TSMLRDPSHI TAETAARRLA
VN004         RDEVSFSVGL SSYAVGSQLP CEPEPDVTVV TSMLIDPSHV TAEAAARRLA
VN235         RDDITFSVGL STYVVGSQLP CEPEPDVVIL TSMLTDPDHI TAETAARRLA
VN405         RDEISFLVGL NSYAIGSQLP CEPEPDVTVV TSMLVDPSHL TAEAAARRLA

              2201                                              2250
BEBE1         RGSPPSAASS SASQLSAPSL RATCTTHAK. ...CPDIDMV DANLFCWCTM
D89815        RGSPPSLAGS SASQLSALSL KATCTTHHG. ...APDTDLI EANLLWRQEM
ED43type_4    RGSRPSLASS SASQLSPRLL QATCTAPHD. ...SPGTDLL EANLLW....
HC_C2         RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
HC_G9         RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI TANLLWRQEM
HCU16326      RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
HCV_H_CMR     RGSPPSMASS SASQLSAPSL KATCTANHD. ...SPDAELI EANLLWRQEM
HCV_J1        RGSPPSEASS SASQLSAPSL KATCTINHD. ...SPDAELI EANLLWRQEM
HCV_J483      RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
HCV_J8        RGSPPSQASS SASQLSAPSL KATCTTHKT. ...AYDCDMV DANLF....M
HCV_JK1       RGSPPSLASS SASQLSAPSL KATCTTRHD. ...SPDADLI EANLLWRQEM
HCV_JS        RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
HCV_K1_R1     RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
HCV_K1_R2     RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
HCV_K1_R3     RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
HCV_K1_S1     RGSPPSLASS SASQLSAPSS KATYITQYD. ...SPDFDLI EANLLWRQEM
HCV_K1_S2     RGSPPSLASS SASQLSAPSL KATCTTRHD. ...SPDADLI EANLLWRQEM
HCV_K1_S3     RGSPPSLASS SASQLSAPSL KATCTTCHD. ...SPDADLI EANLLWRQEM
HCV_L2        RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
HCV_N         RGSPPSLASS SASQLSAPSL RATCTTHSSY NLDSPDVDLI EANLLWRQEM
HCV12083      KGSPPSLASS SANQLSAPSL RATCTTSQK. ...HPEMELL QANLLWKHEM
HCV1480       RGSPPSLANS SASQLSAPSL KATCTIQGH. ...HPDADLI KANLLWRQCM
HCVPOLYP      RGSPPSLASS SASQLSAPSL KATCTTRHD. ...SPDADLI EANLLWRQEM
HD_1          RGSPPSLASS SASQLSAPSL KATCTTRHD. ...SPDADLI EAHLLWRQEM
HPCCGAA       RGSPPSMASS SASQLSAPSL KATCTANHD. ...SPDAELI EANLLWRQEM
HPCFG         RGSPPSEASS SASQLSAPSL KATCQTHRP. ...HPDAELI DANLLWRQEM
HPCGENANTI    RGSPPSLASS SASQLSALSL KAACTTRHT. ...PPDADLI EANLLWRQEM
```

```
HPCGENOM  RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
 HPCHUMR  RGSPPSLASS SASQLSAPSL KATCTTHHV. ...SPDADLI EANLLWRQEM
    HPCJ  RGSPPSLASS SASQLSAPSL KATCTIHHD. ...SPDADLI EANLLWRQEM
   HPCJCG RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
 HPCJK046 RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
 HPCJK049 RGSPPSLASS SASQLSAPSR KATCTTHAD. ...HPDAELV EANLLWRQEM
   HPCJTA RGSPPSLASS SASQLSAPSL KATCTTHGR. ...HPDAELI TANLLWRQEM
   HPCJTB RGSPPSLASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
   HPCK3A RGSPPSEASS SASQLSAPSL KATCTTHHD. ...SPDADLI EANLLWRQEM
HPCPLYPRE RGSPPSVASS SASQLSAPSL KATCQTHRP. ...HPDAELV DANLLWRQEM
  HPCPOLP RGSPPSEASS SASQLSAPSL KATCTANHD. ...SPDAELI EANLLWRQEM
    HPCPP RGSPPSLASS SASQLSALSL RATCTTHGK. ...AYDVDMV DANLF....M
 HPCUNKCD RGSPPSLASS SASQLSAPSL KATCTTHHG. ...APDTDLI EANLLWRQEM
   MKC1A  RGSPPSLASS SASQLSALSL KATCTTHHD. ...SPDADLI EANLLWRQEM
   NDM59  RGSPPSEASS SASQLSAPSL KATCTTHHG. ...APDTDLI EANLLWRQEM
    NZLI  RGSPPSEASS SASQLSAPSL RATCTTHGK. ...AYDVDMV DANLF....M
    SA13  RGSPPSLASS SASQLSAPSL KATCQTHRP. ...HPDAELV DANLLWRQEM
   Th580  RGSPPSLASS SASQLSAPSL KATCTTQGH. ...HPDADLI EANLLWRQCM
Type_3a_CB RGSPPSEASS SASQLSAPSL KATCTANGD. ...HPDAELI EANLLWRQEM
  TypeV_D RGSPPSEASS SASQLSAPSL KATCQTHRP. ...HPDAELV NANLLWRQEM
   VN004  RGSPPSLASS SASQLSAPSL KATCTMHGA. ...HPDAELI EANLLWRQEM
   VN235  RGSPPSLASS SASQLSAPSL KATCTTAGK. ...HPDAELI EANLLWRQEV
   VN405  RGSPPSCASS LASQLSAPSL KATCTTHCA. ...HPDADLI EANLLWRQEV

          2251                                                2300
    BEBE1 GGNMTRIESE SKVLMVDSFD PVVDKE.DER EPSIPSEYLL PKS.RFPPAL
   D89815 GGNITRVESE NKIVILDSFE PLRAEE.DER EVSAAAEILR KTR.KFPAAM
ED43type_4 GSTATRVETD EKVIILDSFE SCVAEQNDDR EVSVAAEILR PTK.KFPPAL
    HC_C2 GGNITRVESE NKVVILDSFE PLRAEE.DER EVSVAAEILR PTK.KFPPAL
    HC_G9 GGNITRVESE NKIVILDSFD PLVAEE.DDR EISVPAEILL KSK.KFPPAM
  HCU16326 GGNITRVESE NKVVILDSFD PLRAED.DEG EISVPAEILR KSR.KFPPAL
 HCV_H_CMR GGNITRVESE NKVVILDSFD PLVAEE.DER EVSVPAEILR KSR.RFARAL
    HCV_J1 GGNITRVESE NKVVILDSFD PLVAEE.DER EISVPAEILR KSR.RFTQAL
  HCV_J483 GGNITRVESE NKVVILDSFE PLHAEG.DER EISVAAEILR KSR.KFPSAL
    HCV_J8 GGDVTRIESD SKVIVLDSLD SMTEVE.DDR EPSVPSEYLI KRR.KFPPAL
   HCV_JK1 GGNITRVESE NKVVILDSFE PLRAEE.DER EVSVAAEILR KSR.KFPPAL
    HCV_JS GGNITRVESE NKVVILDSFD PLHAEE.DER EVSVAAEILR KSR.KFPPAL
  HCV_K1_R1 GGNITRVESE NKVIILDSFD PLRAEE.DER EVSIPAEILR KSK.KFPPAL
  HCV_K1_R2 GGNITRVESE NKVVILDSFE PLRAEE.DER EVSLPAEILR KSR.KFPRAM
  HCV_K1_R3 GGNITRVESE NKVVILDSFD PLRAEE.DER EVSVAAEILR KTR.KFPPAL
  HCV_K1_S1 GGNITRVESE NKVVTLDSFD PLRAEE.DER EVSIPAEILR KSK.KFPSAL
  HCV_K1_S2 GGNITRVESE NKVVILDSFE PLRAEE.DER EVSLPAEILR KSR.KFPPAM
  HCV_K1_S3 GGNITRVESE NKVVILDSFD PLRAEE.DER EVSVAAEILR KTK.KFPPAL
    HCV_L2 GGNITRVESE SKVVILDSFD PLRAEE.GEG EVSVAAEILR KSK.KFPPAL
     HCV_N GGNITRVESE NKVVVLDSFE PLRAEG.DEN EISIAAEILR KSK.KFPAAI
  HCV12083 GSHIPRVQSE NKVVVLDSFE LYPLEY.EER EISVSVECHR QPRCKFPPVF
  HCV1480 GGNITRVEAE NKVEILDCFK PLKEEE.DDR EISVSADCFK KGP.AFPPAL
 HCVPOLYP GGNITRVESE NKVVILDSFD PLRAEE.DER EVSVAAEILR KSR.RFPRAM
     HD_1 GGNITRVESE NKVVILDSFD PLRAEE.DER EVSVPAEILR KSR.KFPPAM
   HPCCGAA GGNITRVESE NKVVILDSFD PLVAEE.DER EVSVPAEILR KSR.RFAPAL
    HPCFG GSNITRVESE TKVVILDSFE PLRAEE.DDT ELSIPAECFK KPP.KYPPAL
HPCGENANTI GGNITRVESE NKVVILDSFD PLRAEE.DER EVSVPAEILR KSR.KFPPAL
 HPCGENOM GGNITRVESE NKVVILDSFD PLRAEE.DER EVSVAAEILR KSR.KFPSAL
  HPCHUMR GGNITRVESE NKVVVLDSFD PLRAEE.DER EVSVPAEILR KSK.KFPAAM
    HPCJ  GGNITRVESE NKVVILDSFE PIRAEE.DER EVSVPAEILR RSR.KFPAAM
   HPCJCG GGNITRVESE NKVVILDSFD PIRAVE.DER EISVPAEILR KPR.KFPPAL
 HPCJK046 GGNITRVESE NKIVILDSFE PLKAEF.DDR EISVAAECHR PPRFKYPPAL
 HPCJK049 GSNITRVESE SKVVILDSFE PLRACD.DED ELSVAAECFK KPP.KYPPAL
   HPCJTA GGNITRVESE NKVVILDSFD PLRAEE.DER EVSVAAEILR KSK.KFPPAL
   HPCJTB GGNITRVESE NKVVILDSFD PLRAEE.DER EVSVAAEILR KSK.KFPPAL
   HPCK3A GSNITRVESE TKVVILDSFE PLRAET.DDA ELSAAAECFK KPP.KYPPAL
```

```
HPCPLYPRE  GGNITRVESE NKVVILDSFD PLVAEE.DER EISVPAEILR KSR.RFAQAL
  HPCPOLP  GGDVTRIESE SKVVVLDSLD PMVEER.SDL EPSIPSEYML PKK.RFPPAL
    HPCPP  GGNITRVESE NKIVILDSFE PLRAEE.DER EVSVAAEILR KTR.KFPAAM
 HPCUNKCD  GGNITRVESE NKVVILDSFD PLRAED.DEG EISVPAEILR KSR.KFPPAL
    MKC1A  GGNITRVESE NKIVILDSFE PLRAEE.DER EVSVAAEILR KTR.KFPAAM
    NDM59  GGDVTRIESE SKVVVLDSLD PMAEER.SDL EPSIPSEYML PRN.RFPPAL
     NZLI  GSNITRVESE TKVVVLDSFE PLRAET.DDV EPSVAAECFK KPP.KYPPAL
     SA13  GGNITRVEAE NKVVILDSFE PLKADD.DDR EISVSADCFR RGP.AFPPAL
     Th580 GSNITRVESE TKVVILDSFD PLVAEY.DDR EISVSAECHR PPRPKFPPAL
Type_3a_CB GSNITRVESE TKVVILDSFE PLRAET.DDA ELSVAAECFK KPP.KYPPAL
  TypeV_D  GSNITRVESE TKVVILDSFE PLRAQT.DDA ELSVAAECFK KPP.KYPPAL
    VN004  GGNITRVESE NKVVILDSFD PLVPEF.EER EMSVPAECHR PRRPKFPPAL
    VN235  GGNITRVESE NKIIVLDSFD PLIAET.DDR EISVGAECFN PPRPKFPPAL
    VN405  GGNITRVESE NKVIVLDSFD PLVPEY.DDR EPSVPAECHR PNRPKFPPAL

           2301                                                2350
    BEBE1  PPWARPDYNP PLLETWKRPD YQPPVVAGCA LPPPGTTPVP PPRRRR.AVV
    D89815 PVWARPDYNP PLLESWKNPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
ED43type_4 PIWARPDYNP PLTETWKQQD YQAPTVHGCA LPPAKQPPVP SPRRKR.TVQ
    HC_C2  PVWARPDYNP PLLESWKDPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
    HC_G9  PIWARPDYNP PLVEPWKRPD YEPPLVHGCP LPPPKPTPVP PPRRKR.TVV
 HCU16326  PIWAPPDYNP PLLESWKDPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
 HCV_H_CMR PVWARPDYNP PLVETWKKPD YEPPVVHGCP LPPPRSPPVP PPRKKR.TVV
    HCV_J1 PIWARPDYNP PLIETWKKPN YEPPVVHGCP LPPPQSPPVP PPRKKR.TVV
  HCV_J483 PIWARPDYNP PLLESWKDPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
    HCV_J8 PPWARPDYNP VLIETWKRPG YEPPTVLGCA LPPTPQTPVP PPRRRR.AKV
   HCV_JK1 PIWARPSYNP PLLESWKDPD YVPPVVHGCP LPPTMAPPIP PPRRKR.TVV
    HCV_JS PIWARPDYNP PLLESWKDPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
 HCV_K1_R1 PIWARPDYNP PLLESWKDPD YVPPVVHGCP LPPAKAPPIP PPRRKR.TVV
 HCV_K1_R2 PIWARPDYNP PLIESWKDPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
 HCV_K1_R3 PIWARPDYNP PLLEPWRDPD YAPPVVHGCP LPPAKDPPIP PPRRKR.TVV
 HCV_K1_S1 PIWARPDYNP PLLEPWKDPD YVPPVVHGCP LPPAKAPPIP PPRRKR.TVV
 HCV_K1_S2 PIWARPDYNP PLIESWKDPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
 HCV_K1_S3 PIWARPDYNP PLLESWRAPD YAPPVVHGCP LPPAKDPPIP PPRRKR.TVV
    HCV_L2 PEWARPDYNP PLLESWKDPD YVPPVVHGCP LPPAKAPPIP PPRRKR.TVV
     HCV_N PIWARPDYNP PLLESWKNPD YVPPVVHGCP LPPVKAPPIP PPRRKR.TVV
  HCV12083 PVWARPDNNP PFIQAWQMPG YEPPVVSGCA VAPPKPAPVP PPRRKR.LVH
   HCV1480 PVWARPGYDP PLLETWKRPD YDPPQVWGCP IPPAGPPPVP LPRRKRKPME
  HCVPOLYP PIWARPDYNP PLLESWKDPD YVPPVVHGCP LPPTKAPPIP PPRKKR.TVV
     HD_1  PIWARPDYNP PLLESWKDPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
   HPCCGAA PVWARPDYNP LLVETWKKPD YEPPVVHGCP LPPPRSPPVP PPRKKR.TVV
    HPCFG  PIWARPDYNP PLLPSWKDPT YEPPAVHGCA LPPTRPAPVP PPRRKR.TIK
HPCGENANTI PVWARPDYNP PLLEPWKDPD YVPPVVHGCP LPPVKAPPIP PPRRKR.TVV
  HPCGENOM PIWARPDYNP PLLESWKDPD YVPPVVHGCP LPPTTAPPVP PPRRKR.TVV
   HPCHUMR PIWARPDYNP PLLESWKDPD YVPPVVHGCP LPPIKAPPIP PPRRKR.TVV
     HPCJ  PIWARPDYNP PLLESWKDPD YVPPVVHGCP LPPTKAAPIP PPRRKR.TIV
    HPCJCG PIWARPDYNP PLLESWKDPD YVPPVVHGCP LPSTKAPPIP PPRRKR.TVV
  HPCJK046 PVWARPDYNP PLLETWKAPD YDPPVVSGCA LPPQGLPPVP PPRRKK.LVQ
  HPCJK049 PIWARPDYNP PLVEPWKDPD YVPPTVHGCA LPPQKLPPVP PPRRKR.TIV
    HPCJTA PIWARPDYNP PLLESWKSPD YVPPAVHGCP LPPTTGPPIP PPRKKR.TVV
    HPCJTB PIWARPDYNP PLLESWKSPD YVPPAVHGCP LPPTTGPPIP PPRKKR.TVV
    HPCK3A PIWARPDYNP PLLDRWKSPD YVPPTVHGCA LPPKGAPPVP PPRRKR.TIQ
 HPCPLYPRE PVWARPDYNP PLVETWKKPD YEPPVVHGCP LPPPKSPPVP PPRKKR.TVV
   HPCPOLP PAWARPDYNP PLVESWKRPD YQPATVAGCA LPPPKKTPTP PPRRRR.TVG
    HPCPP  PVWARPDYNP PLLESWKNPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
 HPCUNKCD  PIWAPPDYNP PLLESWKDPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
    MKC1A  PVWARPDYNP PLLESWKNPD YVPPVVHGCP LPPTKAPPIP PPRRKR.TVV
    NDM59  PAWARPDYNP PLVESWKRPD YQPPTVAGCA LPPPKKTPTP PPRRRR.TVG
     NZLI  PIWARPDYNP PLLDRWKAPD YVPPTVHGCA LPPRGAPPVP PPRRKR.TIQ
     SA13  PIWARPGYDP PLLETWKQPD YDPPQVSGCP LPPAGLPPVP PPRRKRKPVV
     Th580 PIWARPDYNP PLLQKWQMPG YEPPVVSGCA LPPAKPTPIP PPRRKR.LIQ
```

```
Type_3a_CB  PIWARPDYNP PLLDRWKAPD YVPPTVHGCA LPPRGAPPVP PPRRKR.TIQ
   TypeV_D  PIWARPDYNP PLLDRWKTPD YVPPTVHGCA LPPRGAPPVP PPRRKR.TIQ
      VN004 PIWATPGYNP PVLETWKSPT YEPPVVHGCA LPPSGPPPIP PPRRKK.VVQ
      VN235 PVWARPDYNP PLLQPWKAPD YEPPLVHGCA LPPKGLPPVP PPRKKR.VVQ
      VN405 PIWARPDYNP PLLETWKKPD YAPPLVHGCA LPSPVQPPVP PPRRKS.VVH

            2351                                                2400
     BEBE1  LDQSNVGEAL KELAIKSFGC PPPSGDPGHS TGGGTTGETS KSPPD.EPDD
    D89815  LTESTVSSAL AELATKTFGG SGS.SAVDSG TATGPPDQAS AE....GDAG
ED43type_4  LTESVVSTAL AELAAKTFGQ SEP.SSDRDT DLTTPTETTD SGPIV.VDDA
     HC_C2  LTESTVSSAL AELATKTFGS SGS.SAVDSG TATAPPDQTS ND....GDRE
     HC_G9  LDESTVSSAL AELATKTFGS STT.SGVTSG EAAESSPAPS CD....GELD
   HCU16326 LTESTVSSAL AELATKTFGS SGS.SAIDSG TATAPPDQAS GD....GDRE
  HCV_H_CMR LTESTLPTAL AELATKSFGS SST.SGITGD NMTTSSEPAP SG....CPPD
     HCV_J1 LTESTLSTAL AELAAKSFGS SST.SGITGD NTTTSSEPAP SG....CSPD
   HCV_J483 LTESNVSSAL AELATKTFSS SGS.SAVDSG TATALPDQAS DD....GDKG
     HCV_J8 LTQDNVEGVL REMADKVLSP LQDNNDSGHS TGADTGGDIV QQPSD.ETAA
    HCV_JK1 LTESTVSSAL AELATKTFGS SGS.SAVDSG TATAPPDQPS DD....GDRG
     HCV_JS LTESTVSSAL AELATKTFGS SGS.SAADSG TATAPPDQAS DD....GDKG
  HCV_K1_R1 LTESTVSSAL AELATKTFGS SGS.SAADRG TATAPPDQAS ND....GDAG
  HCV_K1_R2 LTESTVSSAL AELATKTFGS SES.SAADSG TATAPPDQPS SD....GDAG
  HCV_K1_R3 LTESTVSSAL AELATKTFGS SGS.SAVDSG TATAPPDQTS ND....GDTG
  HCV_K1_S1 LTESTVSSAL AELATKTFGS SES.SAADRG TATAPPDQTS ND....GDAG
  HCV_K1_S2 LTESTVSSAL AELATKTFGS SES.SAADSG TATAPPDQPS ND....GDAG
  HCV_K1_S3 LTESTVSSAL AELATKTFGS SGS.SAVDSG TATAPPDQPS ND....GDTG
     HCV_L2 LTESTVSSAL AELAVKTFGS SES.SAVDSG TATAPPDQVS DN....GDKG
      HCV_N LTDSTVSSVL AELATKTFGS SEL.SAADSG TATAPPDQTS DN....GGKD
   HCV12083 LDESTVSHAL AQLADKVFVE SSNDPGPSSD SGLSITSPVP PDPTTPEDAG
    HCV1480 LSDSTVSQVM ADLADARFKV DTP.SIEGQD SALGTSSQHD SGPEEKRDDN
   HCVPOLYP LTESTVSSAL AELATKTFGS SES.SAVDSG TATAPPDQPP DN....DDTG
       HD_1 LTESTVSSAL AELATKTFGS SES.SAVDSG TATAPPGQSS DD....VDTG
    HPCCGAA LTESTLPTAL AELATKSFGS SST.SGITGD NTTTSSEPAP SG....CPPD
      HPCFG LDGSNVSAAL LALAERSFPS TKPEGTGTSS SGVGTESTAE SGDSPETGEE
 HPCGENANTI LTESTVSSAL AELATKTFGS SES.SAAGSG TATAPPDQPS DD....GDAG
   HPCGENOM LTESSVSSAL AELATKTFGS SES.SAVDSG TATAPPDEAS GG....GDKG
   HPCHUMR LTESSVSSAL AELATKTFGS SES.SAVDSG TATALPDQAS DD....GDKG
      HPCJ LTESTVSSAL AELATKTFGG SGS.SAADSG TATAPPDQTS DD....GDKE
    HPCJCG LTESTVSSAL AELATKTFGS SGS.SAVDSG TATGPPDQAS DD....GDKG
   HPCJK046 LDDSVVGHVL AQLAEKSFPA TPDQPQTNSD SGHGTNGAAS LPSAE.DDDA
   HPCJK049 LSESTVSKAL ASLAEKSFPQ PTCSAEDEST SGVGTQSGSL TGPVQLDDDD
    HPCJTA LTESTVSSAL AELATKTFGS SGS.SAVDSG TATAPPDQTS DD....GDKE
    HPCJTB LTESTVSSAL AELATKTFGS SGS.SAVDSG TATAPPDQTS DD....GDKE
    HPCK3A LDGSNVSAAL AALAEKSFPS SKPQEENSSS SGVDTQSSTA SKVLPSPGEE
 HPCPLYPRE LTESTLSTAL AELATRSFGS SST.SGITGD NTTTSSEPAP SG....CPPD
   HPCPOLP LSESSIADAL QQLAIKSFGQ PPPSGDSGLS TGADAADSGS RTPPD.ELAL
     HPCPP LTESTVSSAL AELATKTFGS SGS.SAVDSG TATGPPDQAS AE....GDAG
  HPCUNKCD LTESTVSSAL AELATKTFGS SGS.SAIDSG TATAPPDQAS GD....GDRE
     MKC1A LTESTVSSAL AELATKTFGS SGS.SAVDSG TATGPPDQAS AE....GDAG
     NDM59 LNENTIGDAL QQLAIKAFGQ PPLSGDSGLS TGADAADSGS RTPPD.ESAL
      NZLI LDGSNVSAAL AALAEKSFPS SKPQEENSSS SGVDTQSSTT SKVPPSPGGE
      SA13 LSDSNVSQVL ADLAHARFKA DTQ.SIEGQD SAVGTSSQPD SGPEEKRDDD
     Th580 LDESAVSQAL QQLADKVFVE DTSTSEPSSG LGGSIAGPSS PDPTTADDTC
Type_3a_CB LDGSNVSAAL RALAEKSFPS LKPQEENNSS SGVDTQSSTT SKVPPSPGGE
   TypeV_D LDGSNVSAAL AALAKKSFPS VNPQDENSSS SGVDTQSSTT SKVPPSPGGE
     VN004 LDSSNVSAAL AQLAAKTFET PSS.PTTGYG SDQPDHSTES SEHDRDDGVA
     VN235 LDEGSAKRAL AELAQTSFPP STATLSEDSG RETSTLSSDM TPPREEADRA
     VN405 LDDSTVATAL AELAEKSFPT QPA.STPDSD SGHPTTSKSS DQADEGEDTP

            2401                                                2450
     BEBE1  SEAGSVSSMP PLEGEPGDPD LEPEQVEHPA PPQEGGAAPG SDSGSWSTCS
    D89815  SDAESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
```

```
ED43type_4  SDDGSYSSMP PLEGEPGDPD LT..... ..... ..SDSWSTVS
      HC_C2 SDAESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
      HC_G9 SEAESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
    HCU16326 SDVESFSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
   HCV_H_CMR SDVESYSSMP PLEGEPGDPD FS..... ..... ..DGSWSTVS
     HCV_J1 SDAESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
    HCV_J483 SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
     HCV_J8 SEAGSLSSMP PLEGEPGDPD LEFEPVGSAP PSEGECEVID SDSKSWSTVS
    HCV_JK1 SDDESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
     HCV_JS SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
   HCV_K1_R1 SDVGSYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
   HCV_K1_R2 SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
   HCV_K1_R3 SDVGSYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
   HCV_K1_S1 SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
   HCV_K1_S2 SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
   HCV_K1_S3 SDVGSYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
     HCV_L2 SDAESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
      HCV_N SDAESCSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
    HCV12083 SEAESYSSMP PLEGEPGDPD LS..... ..... ..SGSWSTVS
     HCV1480 SDAASYSSMP PLEGEPGDPD LS..... ..... ..SGSWSTVS
    HCVPOLYP SDVESCSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
        HD_1 SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
     HPCCGAA SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
      HPCFG SDVESYSSMP PLEGEPGDPD LD..... ..... ..ADSWSTVS
  HPCGENANTI SDVESCSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
    HPCGENOM SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
     HPCHUMR SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
        HPCJ SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
      HPCJCG SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
    HPCJK046 SDADSYSSMP PLEGEPGDPD LSDG..... ..... GGSGSWSTVS
    HPCJK049 SDNESHSSMP PLEGEPGDPD LS..... ..... ..SGSWSTVS
      HPCJTA SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
      HPCJTB SDVESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
      HPCK3A SDSESCSSMP PLEGEPGDPD LS..... ..... ..CDSWSTVS
   HPCPLYPRE SDAESYSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
     HPCPOLP SETGSISSMP PLEGEPGDPD LEPEQVELQP PPQGGVVTPG SGSGSWSTCS
      HPCPP SDAESYSSMP PLEGEPGDPD LN..... ..... ..DGSWSTVS
    HPCUNKCD SDVESFSSMP PLEGEPGDPD LS..... ..... ..DGSWSTVS
      MKC1A SDAESYSSMP PLEGEPGDPD LN..... ..... ..DGSWSTVS
      NDM59 SETGSISSMP PLEGEPGDPD LEPEQVELQP PPQGEEVVPG SDSGSWSTCS
       NZLI SDSESCSSMP PLEGEPGDPD LS..... ..... ..CDSWSTVS
       SA13 SDAASYSSMP PLEGEPGDPD LS..... ..... ..SGSWSTVS
      Th580 SDAGSFSSMP PLEGEPGDPD LS..... ..... ..TGSWSTVS
  Type_3a_CB SDSESCSSMP PLEGEPGDPD LS..... ..... ..CDSWSTVS
     TypeV_D SDSESCSSMP PLEGEPGDPD LS..... ..... ..CDSWSTVS
      VN004 SEAESYSSMP PLEGEPGDPD LS..... ..... ..SGSWSTVS
      VN235 SDDGSYSSMP PLEGEPGDPD LS..... ..... ..SGSWSTVS
      VN405 SEAGSYSSMP PLEGEPGDPD LS..... ..... ..SGSWSTVS

            2451                                        2500
      BEBE1 DVD..DSVVC CSMSYSWTGA LITPCSPEEE KLPINPLSNS LLRYHNKVYC
     D89815 EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNP LLRHHNMVYS
ED43type_4 GSE...DVVC CSMSYSWTGA LVTPCAAEES KLPISPLSNS LLRHHNMVYA
      HC_C2 EEA.SGDVVC CSMSYTWTGA LITPCAAEES KLPINPLSNS LLRHHNMVYA
      HC_G9 SDGGTEDVVC CSMSYSWTGA LITPCAAEET KLPINALSNS LLRHHNLVYS
    HCU16326 EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINPLSNS LLRHHNMVYA
   HCV_H_CMR SGADTEDVVC CSMSYTWTGA LVTPCAAEEQ KLPINALSNS LLRHHNLVYS
     HCV_J1 SEAGTEDVVC CSMSYTWTGA LITPCAAEEQ KLPINALSNS LLRHHNLVYS
    HCV_J483 EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINPLSNS LLRHHNMVYA
     HCV_J8 DQE..DSVIC CSMSYSWTGA LITPCGPEEE KLPINPLSNS LMRFHNKVYS
    HCV_JK1 EEA.SEDVAC CSMSYTWTGA LITPCAAEES KLPINPLSNS LLRHHNMVYA
```

```
HCV_JS      EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINPLSNS LLRHHNMVYA
HCV_K1_R1   EEA.GEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HCV_K1_R2   EEA.GEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HCV_K1_R3   EEA.GEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HCV_K1_S1   EEA.GEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HCV_K1_S2   EEA.GEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HCV_K1_S3   EEA.GEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HCV_L2      EEA.SEDVVC CSMSYSWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HCV_N       EEA.GESVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HCV12083    DEDD...VVC CSMSYSWTGA LITPCAAEEE KLPINPLSNS LVRHHNMVYS
HCV1480     GED...NVVC CSMSYTWTGA LITPCSAEEE KLPINPLSNT LLRHHNLVYS
HCVPOLYP    EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HD_1        EEA.NEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HPCCGAA     SGADTEDVVC CSMSYSWTGA LVTPCAAEEQ KLPINALSNS LLRHHNLVYS
HPCFG       DSE.EQSVVC CSMSYSWTGA IITPCSAEEE KLPISPLSNS LLRHHNLVYS
HPCGENANTI  EED.GEGVIC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HPCGENOM    EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINPLSNS LLRHHNMVYA
HPCHUMR     EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HPCJ        EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINPLSNS LLRHHNMVYA
HPCJCG      GEA.GEDVVC CSMSYTWTGA LITPCAAEES KLPINPLSNS LLRHHSMVYS
HPCJK046    SEETS..VVC CSMSYSWTGA LITPCAAEEE KLPISPLSNT LIRHHNMVYS
HPCJK049    .GE.EQSVVC CSMSYSWTGA LITPCAAEEE KLPISPLSNS LLRHHNLVYS
HPCJTA      GEA.SDDIVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HPCJTB      GEA.SDDIVC CSMSYTWTGA LITPCAAEES KLPINALSNS LLRHHNMVYA
HPCK3A      DSE.EQSVVC CSMSYSWTGA LITPCSAEEE KLPISPLSNS LLRHHNLVYS
HPCPLYPRE   SEANAEDVVC CSMSYSWTGA LVTPCAAEEQ KLPINALSNS LLRHHNLVYS
HPCPOLP     EED..DSVVC CSMSYSWTGA LITPCSPEEE KLPINPLSNS LLRYHNKVYC
HPCPP       EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNP LLRHHNMVYA
HPCUNKCD    EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINPLSNS LLRHHNMVYA
MKC1A       EEA.SEDVVC CSMSYTWTGA LITPCAAEES KLPINALSNP LLRHHNMVYA
NDM59       EED..DSVVC CSMSYSWTGA LITPCSPEEE KLPINPLSNS LLRYHNKVYC
NZLI        DSE.EQSVVC CSMSYSWTGA LITPCSAEEE KLPISPLSNS LLRHHNLVYS
SA13        DED...SVVC CSMSYSWTGA LITPCSAEEE KLPINPLSNT LLRHHNLVYS
Th580       EEDD...VVC CSMSYTWTGA LITPCAAEEE KLPINPLSNS LIRHHNMVYS
Type_3a_CB  DSE.EQSVVC CSMSYSWTGA LITPCSAEEE KLPISPLSNS LLRHHNLVYS
TypeV_D     DSE.EQSVVC CSMSYSWTGA LITPCSAEEE KLPISPLSNS LLRHHNLVYS
VN004       EEGDS..VVC CSYSYSWTGA LVTPCAAEEE KLPINPLSNS LIRHHNLVYS
VN235       EDHDS..VVC CSMSYSWTGA LITPCAAEEE KLPISPLSNA LIRHHNLVYS
VN405       EEGDS..VVC CSMSYSWTGA LVTPCAAEEE KLPINPLSNS LIRHHNLVYS

            2501                                                2550
BEBE1       TTSRSASQRA KKVTFDRVQL LDSHYESVLK DVKQAATKVS AKLLSIEEAC
D89815      TTSRSASLRQ KKVTFDRMQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
ED43type_4  TTTRSAVTRQ KKVTFDRLQV VDSTYNEVLK EIKARASRVK PRLLTTEEAC
HC_C2       TTSRSASLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSIEEAC
HC_G9       TTSRSAGQRQ KKVTFDRLQV LDDHYRDVLK EAKAKASTVK AKLLSVEEAC
HCU16326    TTSRSAGLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
HCV_H_CMR   TTSRSACQRQ KKVTFDRLQV LDSHYQDVLK EVKAAASKVK ANLLSVEEAC
HCV_J1      TTSRSACQRQ KKVTFDRLQV LDSHYQDVLK EVKAAASKVK ANLLSVEEAC
HCV_J483    TTSRSASLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSIEEAC
HCV_J8      TTSRSASLRA KKVTFDRVQV LDAHYDSVLQ DVKRAASKVS ARLLTVEEAC
HCV_JK1     TTSRSAGLRQ KKVTFDRLQV PDDHYRDVLK EMKAKASTVK AKLLSVEEAC
HCV_JS      TTSRSAGLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSIEEAC
HCV_K1_R1   TTSRSASQRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
HCV_K1_R2   TTSRSASQRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK ARLLSVEEAC
HCV_K1_R3   TTSRSASQRQ RKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
HCV_K1_S1   TTSRSASQRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
HCV_K1_S2   TTSRSASQRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK ARLLSVEEAC
HCV_K1_S3   TTSRSASQRQ RKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
HCV_L2      TTSRSAGLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
HCV_N       TTSRSAGLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSIEEAC
```

```
HCV12083  TTSRSASLRQ KKVTFDRVQV FDQHYQDVLK EIKLRASTVQ AKLLSIEEAC
 HCV1480  TSSRSAGLRQ KKVTFDRLQV LDDHYREVVD EMKRLASKVK ARLLPLEEAC
HCVPOLYP  TTSRSASQRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
    HD_1  TTSRSASQRQ KKVTFDRQQV QDDHYRDVLK EMKAKASTVK ARLLSVEEAC
 HPCCGAA  TTSRSACQRK KKVTFDRLQV LDSHYQDVLK EVKAAASKVK ANLLSVEEAC
   HPCFG  TSSRSAAARQ KKVTFDRLQV LDDHYKNVLK EVKERASGVK GRLLSFEEAC
HPCGENANTI TTSRSASQRQ KKVTIDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
 HPCGENOM  TTSRSASLRQ KKVAFDRMQV LDDHYRDVLK EMKAKASTVK AKLLSIEEAC
 HPCHUMR  TTSRSAGLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
    HPCJ  TTSRSASLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
  HPCJCG  TTSRSASLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK ARLLSIEEAC
 HPCJK046 TTSRSAALRQ KKVTFDRQQV VDQHYYDTLK EMKARASTVS AKLLSVEEAC
 HPCJK049 TSSRSAAQRQ KKVTFDRLQV LDDHYNTTLK EIKELASGVK AELLSVEEAC
  HPCJTA  TTSRSASLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
  HPCJTB  TTSRSASLRQ KKVTFDRLQV LDDHYRDVLK EVKAKASTVK AKLLSVEEAC
  HPCK3A  TSSRSASQRQ KKVTFDRLQV LDDHYKTALQ EVKERASRVK ARMLSIEEAC
HPCPLYPRE TTSRSACQRQ KKVTFDRLQV LDSHYQDVLK EVKAAASKVK ANLLSVEEAC
 HPCPOLP  TTSKSASLRA KKVTFDRMQA LDAHYDSVLK DIKLAASKVT ARLLTLEEAC
   HPCPP  TTSRSASQRQ KKVTFDRLQV LDDHYRDVLK DMKAKASTVK AKLLSVEEAC
HPCUNKCD  TTSRSAGLRQ KKVTFDRLQV LDDHYRDVLK EMKAKASTVK AKLLSVEEAC
   MKC1A  TTSRSASQRQ KKVTFDRLQV LDDHYRDVLK DMKAKASTVK AKLLSVEEAC
   NDM59  TTSKSASLRA KKVTFDRMQV LDAHYDSVLK DIKLAASKVS ARLLTLEEAC
    NZLI  TSSRSASQRQ KKVTFDRLQV LDDHYKTALK EVKERASRVK ARMLTIEEAC
    SA13  TSSRSAGQRQ KKVTFDRLQV LDDHYREVVD EMKRLASKVK ARLLPLEEAC
   Th580  TTSRSAGLRQ KKVTFDRLQV VDQHYQDVLK EIKLRASTVH ARLLSTEEAC
Type_3a_CB TSSRSASQRQ RKVTFDRLQV LDDHYKTVLK EVKERASRVK ARMLTIEEAC
 TypeV_D  TSSRSASQRQ KKVTFDRLQV LDDHYKTALK EVKERASRVK ARMLTIEEAC
   VN004  TSSRSAATRQ KKVTFDRVQL LDQHYYDTVK EIKLRASHVK AQLLSTEEAC
   VN235  TTSRSASLRQ KKVTFDRVQV VDQHYYDVLK EIKTKASGVS AKLLSVEEAC
   VN405  TTTRSAAMRQ KKVTFDRLQI LDQHYNNVVK EVKLRASGVT AKLLSVEEAC

          2551                                              2600
    BEBE1 ALTPPHSARS KYGFGAKEVR SLSRRAVDHI KSVWEDLLED HCSPIDTTIM
   D89815 KLTPPHSAKS KFGYGAKDVR SLSSRAVNHI RSVWKDLLED TDTPIQTTIM
ED43type_4 DLTPPHSARS KFGYGKKDVR SHSRKAINHI SSVWKDLLDD NNTPIPTTIM
   HC_C2  KLTPPHSAKS KFGYGAKDVR NLSSKAVNHI RSVWKDLLED TETPIDTTIM
   HC_G9  SLTPPHSARS KFGYGAKDVR SHSSKAIRHI NSVWQDLLED NTTPIDTTIM
 HCU16326 KLTPPHSAKS KFGYGAKDVR SLSSRAVTHI RSVWKDLLED TETPISTTIM
 HCV_H_CMR SLTPPHSAKS KFGYGAKDVR CHARKAVAHI NSVWKDLLED SVTPIDTIIM
   HCV_J1 SLTPPHSAKS KFGYGAKDVR CHARKAVNHI NSVWKDLLED SVTPIQTTIM
 HCV_J483 KLTPPHSAKS KFGYGAKDVR NLSSRAVNHI RSVWEDLLED TETPIDTTIM
   HCV_J8 ALTPPHSAKS RYGFGAKEVR SLSRRAVNHI RSVWEDLLED QHTPIDTTIM
  HCV_JK1 KLTPPHSARS KFGYGAKDVR NLSSKAVNHI HSVWKDLLED TETPIDTTIM
   HCV_JS KLTPPHSAKS KYGYGAKDVR NLSSRAVNHI RSVWEDLLED TETPIDTTIM
 HCV_K1_R1 RLTPPHSARS KFGYGAKDVR NLSSGAVNHI RSVWKDLLED TETPIDTTIM
 HCV_K1_R2 KLTPPHSARS KFGYGAKDVR NLSSRAINHI RSVWKDLQED TETPIDTTIM
 HCV_K1_R3 KLTPPHSAKS KFGYGGKDVR NLSSKAVNHI RSVWKDLLED TETPIDTTVM
 HCV_K1_S1 RLTPPHSARS KFGYGAKDVR NLSSGAVNHI RSVWKDLLED TETPIDTTIM
 HCV_K1_S2 KLTPPHSARS KFGYGAKDVR NLSSRAINHI RSVWKDLLED TETPIDTTIM
 HCV_K1_S3 KLTPPHSAKS KFGYGGKDVR NLSSKAVNHI RSVWKDLLED TETPIDTTVM
   HCV_L2 KLTPPHSAKS KFGYGAKDVR NLSSRAVNHI RSVWKDLLED TETPIDTTIM
    HCV_N RLTPPHSAKS KFGYGAKDVR NLSSRAINHF RSVWEDLLED TVTPIDTTVM
 HCV12083 DLTPSHSARS KYGYGAQDVR SRASKAVDHI PSVWEGLLED SDTPIPTTIM
  HCV1480 GLTPPHSARS KYGYGAKEVR SLDKKALKHI EGVWQDLLDD SDTPLPTTIM
 HCVPOLYP KLTPPHSARS KFGYGAKDVR NLSSKAVNHI RSVWKDLLED TETPIDTTIM
    HD_1  KLTPPLSARS KFGYGAKDVR NLSSKAVNHI RSVWEDLLED NVTPIDTTIM
  HPCCGAA SLAPPHSAKS KFGYGAKDVR CHARKAVAHI NSVWKDLLED SVTPIDTTIM
   HPCFG  SLVPPHSGRS KYGYSAKDVR SLSSKAMNQI RSVWEDLLED NSTPIPTTIM
HPCGENANTI KLTPPHSARS KFGYGAKDVR NLSGKAINHI RSVWKDLLED TETPIDTTIM
 HPCGENOM  KLTPPHSAKS KFGYGAKDVR NLSSKAVNHI RSVWKDLLED NETPINTTIM
 HPCHUMR  KLTPPHSAKS KFGYGAKDVR NLSSKAVNHI HSVWKDLLED TVTPIDTTIM
```

HPCJ	KLTPPHSAKS	KFGYGAKDVR	SLSSRAVNHI	TSVWKDLLED	TKTPIDTTIM
HPCJCG	KLTPPHSAKS	KFGYGAKDVR	SLSSRAVNHI	RSVWEDLLED	TETPIDTTIM
HPCJK046	DLTPAHSARS	KFGYGAKDVR	GRTSKALNHI	NSVWEDLLED	NVTPIPTTIM
HPCJK049	RLVPSHSARS	KFGYGAKEVR	SLSSKAINHI	NSVWEDLLED	NTTPIPTTIM
HPCJTA	KLTPPHSAKS	KFGYGAKDVR	NLSSKAINHI	RSVWKDLLED	TETPIDTTIM
HPCJTB	KLTPPHSAKS	KFGYGAKDVR	NLSSKAINHI	RSVWKDLLED	TETPIDTTIM
HPCK3A	ALVPPHSARS	KFGYSAKDVR	SLSSKAINQI	RSVWEDLLED	TTTPIPTTIM
HPCPLYPRE	SLTPPHSAKS	KFGYGAKDVR	CHARKAVTHI	NSVWKDLLED	NVTPIDTTIM
HPCPOLP	QLTPPHSARS	KYGFGAKEVR	SLSGRAVNHI	KSVWKDLLED	TQTPIPTTIM
HPCPP	KLTPPHSARS	KFGYGAKDVR	SLSSKAVNHI	RSVWKDLLED	TETPIDTTIM
HPCUNKCD	KLTPPHSAKS	KFGYGAKDVR	SLSSRAVTHI	RSVWKDLLED	TETPISTTIM
MKC1A	KLTPPHSARS	KFGYGAKDVR	SLSSKAVNHI	RSVWKDLLED	TETPIDTTIM
NDM59	QLTPPHSARS	KYGFGAKEVR	SLSGRAVNHI	KSVWKDLLED	SQTPIPTTIM
NZLI	ALVPPHSARS	KFGYSAKDVR	SLSSRAINQI	RSVWEDLLED	TTTPIPTTIM
SA13	GLTPPHSARS	KYGYGAKEVR	SLDKKALNHI	KGVWQDLLDD	SDTPLPTTIM
Th580	SLTPPHSARS	RYGYGARDVR	SHTSKAVKHI	DSVWEDLLED	NATPIPTTIM
Type_3a_CB	ALVPPHSARS	KFGYSAKDVR	SLSSRAINQI	RSVWEDLLED	TTTPIPTTIM
TypeV_D	ALVPPHSARS	KFGYSAKDVR	SLSSRAIDQI	RSVWEDLLED	TTTPIPTTIM
VN004	DLTPPHSARS	KFGYGAKDVR	SHASKAINHI	NSVWADLLED	TQTPIPTTIM
VN235	ALTPPHSARS	KFGYGAKEVR	GLASKAVNHI	NSVWEDLLED	NSTPIPTTIM
VN405	SLTPPHSARS	KFGYGAKDVR	SHTSKAINHI	NSVWEDLLED	NQTPIPTTIM

	2601				2650
BEBE1	AKNEVFCVDP	TKGGKKPARL	IVYPDLGVRV	CEKMALYDIT	QKLPVAVMGQ
D89815	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
ED43type_4	AKNEVFAVNP	AKGGRKPARL	IVYPDLGSRV	CEKRALHDVI	KKTALAVMGA
HC_C2	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HC_G9	AKNEVFCVKP	EKGGRKPARL	IVYPDLGVRV	CEKRALYDVV	KQLPIAVMGT
HCU16326	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HCV_H_CMR	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	SKLPLAVMGS
HCV_J1	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	SKLPPAVMGS
HCV_J483	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HCV_J8	AKNEVFCIDP	TKGGKKPARL	IVYPDLGVRV	CEKMALYDIA	QKLPKAIMGP
HCV_JK1	AKNEVFCVQP	EKGGRKPARL	IVFPELGVRV	CEKMALYDVV	STLPQAVMGS
HCV_JS	AKSEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HCV_K1_R1	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HCV_K1_R2	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGA
HCV_K1_R3	AKNEVFCVQP	EKGGRKAARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HCV_K1_S1	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HCV_K1_S2	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGP
HCV_K1_S3	AKNEVFCVQP	EKGGRKAARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HCV_L2	AKSEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGP
HCV_N	AKNEVFCVQP	EKGGQKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HCV12083	AKNEVFCVDP	SKGGRKPARL	IVYPDLGVRV	CEKMALYDVT	QKLPQAVMGP
HCV1480	AKNEVFAVEP	SKGGKKPARL	IVYPDLGVRV	CEKRALYDVA	QKLPTALMGP
HCVPOLYP	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGP
HD_1	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPHTVMGS
HPCCGAA	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	SKLPLAVMGS
HPCFG	AKNEVFSVNP	AKGGRKPARL	IVYPDLGVRV	CEKRALYDVI	QKLSIATMGP
HPCGENANTI	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HPCGENOM	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQPVMGS
HPCHUMR	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQVVMGS
HPCJ	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HPCJCG	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGP
HPCJK046	AKNEVFCVDV	SKGGRKPARL	IVYPDLSVRV	CEKRALYDVT	RKLPVAVMGA
HPCJK049	AKNEVFAVAP	HKGGRKPARL	IVYPDLGVRI	CEKRALYDVI	QKLPSAIMGS
HPCJTA	AKSEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HPCJTB	AKSEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	STLPQAVMGS
HPCK3A	AKNEVFCVDP	AKGGRKAARL	IVYPDLGVRV	CEKRALYDVI	QRLSIETMGS
HPCPLYPRE	AKNEVFCVQP	EKGGRKPARL	IVFPDLGVRV	CEKMALYDVV	TKLPLAVMGS
HPCPOLP	AKNEVFCVDP	TKGGKKAARL	IVYPDLGVRV	CEKMALYDIT	QKLPQAVMGA

```
          HPCPP  AKNEVFCVQP EKGGRKPARL IVFPDLGVRV CEKMALYDVV STLPQAVMGS
        HPCUNKCD AKNEVFCVQP EKGGRKPARL IVFPDLGVRV CEKMALYDVV STLPQAVMGS
          MKC1A  AKNEVFCVQP EKGGRKPARL IVFPDLGVRV CEKMALYDVV STLPQAVMGS
          NDM59  AKNEVFCVDP AKGGKKAARL IVYPDLGVRV CEKMALYDVT QKLPQAVMGA
           NZLI  AKNEVFCVDP AKGGRKPARL IVYPDLGVRV CEKRALYDVI QKLSIETMGP
           SA13  AKNEVFAVEP SKGGKKPARL IVYPDLGVRV CEKRALYDIA QKLPTALMGP
           Th580 AKNEVFCVDP SKGGRKPARL IVYPDLSVRV CEKMALYDVT QKLPKTVMGS
       Type_3a_CB AKNEVFCVDP AKGGRKPARL IVYPDLGVRV CEKRALYDVI QKLAIETIGS
          TypeV_D AKNEVFCVDP ARGGRKRARL IVYPDLGVRV CEKRVLYDVI QKLSIETMGT
           VN004  AKNEVFCVDA SKGGRKSARL IVYPDLGVRV CEKRALFDVT RKLPTAIMGD
           VN235  AKNEVFCVDA QKGGRKPARL IVYPDLGVRV CEKRALYDVT QKLPIAVMGA
           VN405  AKNEVFCADV SKGGRKPARL IVYPDLGVRV CEKRALYDVT RKLPTAIMGD

                  2651                                                2700
           BEBE1  SYGFQYSPAQ RVDFLLQAWK EKKTPMGFSY DTRCFDSTVT ERDIRTEESI
           D89815 SYGFQYSPKQ RVEFLVNTWK AKKCPMGFSY DTRCFDSTVT ENDIRVEESI
      ED43type_4  AYGFQYSPAQ RVEFLLTAWK SKNDPMGFSY DTRCFDSTVT EKDIRVEEEV
           HC_C2  SYGFQYSPGQ RVEFLVNTWK SKKCPMGFSY DTRCFDSTVT ENDIRIEESI
           HC_G9  SYGFQYSPAQ RVDFLLNAWK SKKNPMGFSY DTRCFDSTVT EADIRTEEDL
         HCU16326 SYGFQYSPKQ RVEFLVNTWK SKKCPMGFSY DTRCFDSTVT ENDIRVEESI
         HCV_H_CMR SYGFQYSPGQ RVEFLVQAWK SKKTPMGFPY DTRCFDSTVT ESDIRTEEAI
           HCV_J1 SYGFQYSPGQ RVEFLVQAWK SKRTPMGFSY DTRCFDSTVT ESDIRTEEAI
         HCV_J483 SYGFQYSPKQ RVEFLVNTWK SKKCPMGFSY DTRCFDSTVT ESDIRVEESI
           HCV_J8 SYGFQYSPAE RVDFLLKAWG SKKDPMGFSY DTRCFDSTVT ERDIRTEESI
          HCV_JK1 SYGFQYSPGQ RVEFLVNAWK SKKNPMGFAY CTRCFDSTVT ESDIRVEESI
           HCV_JS SYGFQYSPKQ RVEFLVNTWK SKKCPMGFSY DTRCFDSTVT ENDIRVEESI
         HCV_K1_R1 SYGFQYSPGQ RVEFLVNAWK SKKCPMGFAY DTRCFDSTVT ESDIRVEESI
         HCV_K1_R2 SYGFQYSPGQ RVEFLVNAWK SKKCPMGFAY DTRCFDSTVT ENDIRVEESI
         HCV_K1_R3 SYGFQYSPGQ RVEFLVNAWK SKKSPMGFAY DTRCFDSTVT ENDIRTEESI
         HCV_K1_S1 SYGFQYSPGQ RVEFLVNAWK SKKCPMGFAY DTRCFDSTVT ESDIRVEESI
         HCV_K1_S2 SYGFQYSPGQ RVEFLVNAWK SKKCPMGFAY DTRCFDSTVT ESDIRVEESI
         HCV_K1_S3 SYGFQYSPGQ RVEFLVNAWK SKKSPMGFAY DTRCFDSTVT ENDIRVEESI
           HCV_L2 SYGFQYSPGQ RVEFLVNAWK SKKCPMGFSY DTRCFDSTVT ESDIRTEESI
            HCV_N SYGFQYSPGQ RVEFLVKAWK SKKNPMGFSY DTRCFDSTVT ENDIRVEESI
         HCV12083 AYGFQYSPNQ RVEYLLKMWR SKKVPMGFSY DTRCFDSTVT ERDIRTENDI
          HCV1480 SYGFQYSPAQ RVDFLLKAWK SKKIPMAFSY DTRCFDSTIT EHDIMTEESI
         HCVPOLYP SYGFQYSPGQ RVEFLVNAWK SKKCPMGFAY DTRCFDSTVT ESDIRVEESI
             HD_1 SYGFQYSPGQ RVEFLVNTWK SKKCPMGFAY DTRCFDSTVT ENDIRVEESI
          HPCCGAA SYGFQYSPGQ RVEFLVQAWK SKKTPMGLSY DTRCFDSTVT ESDIRTEEAI
           HPCFG AYGFQYSPKQ RVEHLLKMWT SKKTPLGFSY DTRCFDSTVT EHDIRTEEGI
       HPCGENANTI SYGFQYSPGQ RVEFLVNAWK SKKCPMGFSY DTRCFDSTVT ESDIRVEESI
         HPCGENOM SYGFQYSPGQ RVEFLLNAWK SKENPMGFSY DTRCFDSTVT QNDIRVEESI
          HPCHUMR SYGFQYSPGQ RVEFLVNTWK SKKNPMGFSY DTRCFDSTVT ENDIRVEESI
             HPCJ SYGFQYSPGQ RVEFLVKTWK SKKCPMGFSY DTRCFDSTVT ENDIRIEESI
           HPCJCG SYGFQYSPGQ RVEFLVNTWK SKKCPMGFSY DTRCFDSTVT ENDIRTEESI
         HPCJK046 AYGFQYSPSQ RVEYLLKIWR SKKTPMGFSY DTRCFDSTVT ERDIRTEESI
         HPCJK049 AYGFQYSPKQ RVEYLLKMWN SKKTPLGFSY DTRCFDSTVT EQDIRVEESI
           HPCJTA SYGFQYSPGQ RVEFLVNAWK SKKSPMGFSY DTRCFDSTVT ESDIRVEESI
           HPCJTB SYGFQISPGQ RVEFLVNAWK SKKSPMGFSY DTRCFDSTVT ESDIRVEESI
           HPCK3A AYGFQYSPRQ RVERLLKMWT SKKTPLGFSY DTRCFDSTVT GQDIRVEEAV
        HPCPLYPRE SYGFQYSPGQ RVEFLVQAWK SKKTPMGFSY DTRCFDSTVT ESDIRTEEAI
          HPCPOLP SYGFQYSPAQ RVEFLLKAWA EKKDPMGFSY DTRCFDSTVT ERDIRTEESI
            HPCPP SYGFQYSPGQ RVEFLVNAWK SKKSPMGFAY DTRCFDSTVT ENDIRTEESI
        HPCUNKCD SYGFQYSPKQ RVEFLVNTWK SKKCPMGFSY DTRCFDSTVT ENDIRVEESI
          MKC1A SYGFQYSPGQ RVEFLVNAWK SKKSPMGFAY DTRCFDSTVT ENDIRTEESI
          NDM59 SYGFQYSPAQ RVEFLLKAWA EKKDPMGFSY DTRCFDSTVT ERDIRTEESI
           NZLI AYGFQYSPQQ RVERLLKMWT SKKTPLGFSY DTRCFDSTVT EQDIRVEEEI
           SA13 SYGFQYSPAQ RVEFLLKTWR SKKTPMAFSY DTRCFDSTVT EHDIMTEESI
           Th580 AYGFQYSPSQ RVEYLLKMWR SKKTPMGFSY DTRCFDSTVT ERDIRTEEDI
       Type_3a_CB AYGFQYSPQQ RVERLLKMWT SKKTPLGFSY DTRCFDSTVT EQDIRVEEEI
          TypeV_D AYGFQYSPQQ RVERLLKMWT SKKTPLGFSY DTRCFDSTVT EQDFRVEEEI
```

```
VN004   AYGFQYSPQQ RVDRLLKMWR SKKTPMGFSY DTRCFDSTVT ERDIRTEQDI
VN235   AYGFQYSPKQ RVDYLLKMWR SKKTPMGFSY DTRCFDSTVT ERDIRTEEDI
VN405   AYGFQYSPKQ RVDQLLKMWR SKKTPMGFSY DTRCFDSTVT EHDIKTERDV

            2701                                                2750
BEBE1   YLSCSLPEEA RTAIHSLTER LYVGGPMTNS KGQSCGYRRC RASGVLTTSM
D89815  YQCCDLAPEA RQAIRSLTER LYIGGPMTNS KGQNCGYRRC RASGVLTTSC
ED43type_4 YQCCDLEPEA RKVITALTDR LYVGGPMHNS KGDLCGYRRC RATGVYTTSF
HC_C2   YQCCDLAPEA KQAIKSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HC_G9   YQSCDLVPEA RAAIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HCU16326 YQCCDLAPEA KLAIKSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HCV_H_CMR YQCCDLDPQA RVAIKSLTER LYVGGPLTNS RGENCGYRRC RASGVLTTSC
HCV_J1  YQCCDLDPQA RVAIRSLTER LYVGGPLTNS RGENCGYRRC RASGVLTTSC
HCV_J483 YQCCDLAPEA RQAIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HCV_J8  YQACSLPQEA RTVIHSLTER LYVGGPMTNS KGQSCGYRRC RASGVFTTSM
HCV_JK1 YQCCDLAPEA RQVIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTNC
HCV_JS  YQCCDLAPEA KLAIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HCV_K1_R1 YQCCDLAPEA RQAIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HCV_K1_R2 YQCCDLAPEA RQAIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HCV_K1_R3 YQCCDLAPEA RQVIRSLTER LYVGGPLTNS KGQNCGYRRC RASGVLTTSC
HCV_K1_S1 YQCCDLAPEA RQAIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HCV_K1_S2 YQCCDLAPEA RQAIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HCV_K1_S3 YQCCDLAPEA RQVIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HCV_L2  YQCCDLAPEA KQAIKSLTER LYIGGPLTNS KGQNCGYRRC RASVVLTTSC
HCV_N   YQCCDLAPEA RQAIKSLTER LYIGGPLTNS KGQSCGYRRC RASGVLTTSC
HCV12083 YQSCQLDPVA RRVVSSLTER LYVGGPMANS KGQSCGYRRC RASGVLPTSM
HCV1480 YQSCDLQPEA RVAIRSLTQR LYCGGPMYNS KGQQCGYRRC RASGVFTTSM
HCVPOLYP YQCCDLAPEA RQAIKSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HD_1    YQCCDLGPEA RQAIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HPCCGAA YQCCDLDPQA RVAIKSLTER LYVGGPLTNS RGENCGYRRC RASRVLTTSC
HPCFG   YQCCDLEPEA RKAISALTER LYIGGPMYNS KGLQCGYRRC RASGVLPTSF
HPCGENANTI YQCCDLAPEA RQAIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HPCGENOM YQCCDLAPEA RRAIKSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HPCHUMR YQCCDLAPEA RQAIKSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HPCJ    YQCCDLAPEA RQVIRSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HPCJCG  YQCCDLAPEA RQAIRSLTER LYVGGPLTNS KGQNCGYRRC RASGVLTTSC
HPCJK046 YQCCELDPVA RKAISSLTER LYVGGPMYNS QGQSCGYRRC RASGVLPTSM
HPCJK049 YQACDLKDEA RRVITSLTER LYCGGPMFNS KGQHCGYRRC RASGVLPTSF
HPCJTA  YQCCDLAPEA RQAIKSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HPCJTB  YQCCDLAPEA RQAIKSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
HPCK3A  YQCCNLEPEP GQAISSLTER LYCGGPMNNS KGAQCGYLRC RASGVLPTSF
HPCPLYPRE YQCCDLDPQA RVAIKSLTER LYVGGPLTNS RGENCGYRRC RASGVLTTSC
HPCPOLP YRACSLPEEA HTAIHSLTER LYVGGPMFNS KGQTCGYRRC RASGVLTTSM
HPCPP   YQCCDLDPEA RQAIRSLTER LYIGGPLTNS KGQNCGYRRC RVSGVLTTSC
HPCUNKCD YQCCDLAPEA KLAIKSLTER LYIGGPLTNS KGQNCGYRRC RASGVLTTSC
MKC1A   YQCCDLDPEA RQAIRSLTER LYIGGPLTNS KGQNCGYRRC RVSGVLTTSC
NDM59   YRACSLPEEA HTAIHSLTER LYVGGPMLNS KGQTCGYRRC RASGVLTTSM
NZLI    YQCCNLEPEA RKVISSLTER LYCGGPMFNS KGAQCGYRRC RASGVLPTSF
SA13    YQSCDLQPEA RAAIRSLTQR LYCGGPMYNS KGQQCGYRRC RASGVFTTSM
Th580   YQSCQLDPTA RKAISSLTER LYCGGPMFNS KGESCGYRRC RASGVLTTSL
Type_3a_CB YQCCNLEPEA RKVISSLTER LYCGGPMFNS KGAQCGDRRC RASGVLPTSF
TypeV_D YQCCNLEPEA RKVISSLTER LYCGGPMFNS KGAQCGYRRC RASGVLPTSF
VN004   YLSCQLDPEA RKVIESLTER LYVGGPMYNS KGQLCGQRRC RASGVLPTSM
VN235   YQCCQLDPVA KKAITSLTER LYCGGPMYNS RGQSCGYRRC RASGVLTTSL
VN405   YLSCKLDPVA RKAIESLTER LYIGGPMYNS RGQLCGTRRC RASGVLTTSL

            2751                                                2800
BEBE1   GNTLTCYVKA KAACNAAGIV APTMLVCGDD LVVISESQGV EEDERNLRVF
D89815  GNTLTCYLKA AAACRAAKLQ DCTMLVCGDD LVVICDSAGT QEDAASLRVF
ED43type_4 GNTLTCYLKA TAAIRAAALR DCTMLVCGDD LVVIAESDGV EEDNRALRAF
HC_C2   GNTLTCYLKA SAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAASLRVF
```

```
HC_G9       GNTITCYLKA SAACRAAKLR DCTMLVCGDD LVVICESAGV QEDAANLRAF
HCU16326    GNTLTCYLKA TAACRAAKLR DCTMLVNGDD LVVICESAGT QEDAASLRVF
HCV_H_CMR   GNTLTCYIKA RAARRAAGLQ DCTMLVCGDD LVVICESAGV QEDAASLRAF
HCV_J1      GNTLTCYIKA RAACRAAGLQ DCTMLVCGDD LVVICESAGV QEDAASLRAF
HCV_J483    GNTLTCYLKA TAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAAALRVF
HCV_J8      GNTMTCYIKA LAACKAAGIV DPVMLVCGDD LVVISESQGN EEDERNLRAF
HCV_JK1     GNTLTCYLKA SAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAASLRVF
HCV_JS      GNTLTCYLKA SAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAASLRVF
HCV_K1_R1   GNTLTCYLKA SAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAASLRVF
HCV_K1_R2   GNTLTCYLKA TAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAASLRVF
HCV_K1_R3   GNTLTCYLKA AAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAASLRVF
HCV_K1_S1   GNTLTCYLKA SAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAASLRVF
HCV_K1_S2   GNTLTCYLKA TAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAASLRVF
HCV_K1_S3   GNTLTCYLKA AAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAASLRVF
HCV_L2      GNTLTCYLKA SAACRAAKLQ DCTMLVNGDD LVVICESAGT QGDAASLRVF
HCV_N       GNTLTCYLKA SAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAANLRAF
HCV12083    GNTLTCYLKA QAACRAANIK DCDMLVCGDD LVVICESAGV QEDTASLRAF
HCV1480     GNTMTCYIKA LASCRAAKLR DCTLLVCGDD LVAICESQGT HEDEASLRAF
HCVPOLYP    GNTLTCYLKA SAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAASLRVF
HD_1        GNTLTCYLKA TAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAANLRVF
HPCCGAA     GNTLTRYIKA RAACRAAGLQ DCTMLVCGDD LVVICESAGV QEDAASLRAF
HPCFG       GNTITCYIKA TAASRAAGLK NPSFLVCGDD LVVISESCGV EEDRTALRAF
HPCGENANTI  GNTLTCYLKA SAACRAAKLQ DCTMLVCGDD LVVICESAGT QEDAASLRVF
HPCGENOM    GNTLTCYLKA SAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAASLRVF
HPCHUMR     GNTLTCYLKA SAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAASLRVF
HPCJ        GNTLTCYLKA SAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAASLRVF
HPCJCG      GNTLTCYLKA TAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAAALRAF
HPCJK046    GNTLTCYLKA MAACKAAGLK NFDMLVCGDD LVVISESLGV SEDASALRAF
HPCJK049    GNTVTCYLKA KAATKAAGIK DPSFLVCGDD LVVIAESAGI DEDKSALRAF
HPCJTA      GNTLTCYLKA TAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAASLRVF
HPCJTB      GNTLTCYLKA TAACRAAKLQ DCTMLVNGDD LVVICESAGT QEDAASLRVF
HPCK3A      GNTITCYIKA TAAARAAGLR NPDFLVCGDD LVVVAESDGV DEDRATLRAF
HPCPLYPRE   GNTLTCYIKA RAACRAAGLQ DCTMLVCGDD LVVICESAGV QEDAASLRAF
HPCPOLP     GNTITCYVKA LAACKAAGII APTMLVCGDD LVVISESQGT EEDERNLRAF
HPCPP       GNTLTCYLKA SAACRAAKLQ DCTMLVCGDD LVVICDSAGT QEDAASLRVF
HPCUNKCD    GNTLTCYLKA TAACRAAKLR DCTMLVNGDD LVVICESAGT QEDAASLRVF
MKC1A       GNTLTCYLKA SAACRAAKLQ DCTMLVCGDD LVVICDSAGT QEDAASLRVF
NDM59       GNTITCYIKA LAACKAAGIV APTMLVCGDD LVVISESQGT EEDERNLRAF
NZLI        GNTITCYIKA TAAAKAANLR NPDFLVCGDD LVVVAESDGV DEDRAALRAF
SA13        GNTMTCYIKA LASCRAAKLR DCTLLVCGDD LVAICESQGT HEDEASLRAF
Th580       GNTLTCYLKA QAACRAANIK NFDMLVCGDD LVVICESAGV QEDVVALRAF
Type_3a_CB  GNTITCYIKA TAAANGAGLR DPDFLVCGDD LVVVAESDGV DEDGAALRAF
TypeV_D     GNTITCYIKA TAAAKAAGLR NPDFLVCGDD LVVVAESDGV DEDRTALRAF
VN004       GNTVTCFLKA TAACRAAGFT DYDMLVCGDD LVVVTESAGV NEDIANLRAF
VN235       GNTLTCYLKA QAACRAAKLK DFDMLVCGDD LVVISESMGV AEDASALRAF
VN405       GNTMTCFIKA EAACRAAGLT NYDMLVCGDD LVVIAESAGV QEDASNLRAF

            2801                                                2850
BEBE1       TEAMTRYSAP PGDPPKAEYD LELITSCSSN VSVALDPRGR RRYYLTRDPT
D89815      TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
ED43type_4  TEAMTRYSAP PGDAPQPAYD LELITSCSSN VSVAHDVTGK KVYYLTRDPE
HC_C2       TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
HC_G9       TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDGAGK RVYYLTRDPE
HCU16326    TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
HCV_H_CMR   TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDGAGK RVYYLTRDPT
HCV_J1      TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDGTGK RVYYLTRDPT
HCV_J483    TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
HCV_J8      TEAMTRYSAP PGDLPRPEYD LELITSCSSN VSVALDSRGR RRYFLTRDPT
HCV_JK1     TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
HCV_JS      TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
HCV_K1_R1   TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
```

```
HCV_K1_R2   TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
HCV_K1_R3   TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
HCV_K1_S1   TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
HCV_K1_S2   TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
HCV_K1_S3   TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
   HCV_L2   TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
    HCV_N   TEAMTRYSAP PGDLPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
  HCV12083  TDAMTRYSAP PGDAPQPTYD LELITSCSSN VSVAHEGNGK KYYYLTRDCT
   HCV1480  TEAMTRYSAP PGDPPVPAYD LELVTSCSSN VSVARDASGN RIYYLTRDPQ
  HCVPOLYP  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
      HD_1  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
   HPCCGAA  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDGAGK RVYYLTRDPT
     HPCFG  TEAMTRYSAP PGDAPQPTYD LELISSCSSN VSVACDGAGK RYYYLTRDPE
 HPCGENANTI TEAMTRYSAP PGDLPQPEYD QELITSCSSN VSVAHDASGK RVYYLTRDPT
  HPCGENOM  TEAMTRYSAP PGDLPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
   HPCHUMR  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
      HPCJ  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
    HPCJCG  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
  HPCJK046  TDAMTRYSAP PGDEPHPEYD LEHITSCSSN VSVAHDHTGQ RYYYLTRDPT
  HPCJK049  TEAMTRYSAP PGDPPQPTYD LELITSCSSN VSVAHDGAGK RYYYLTRDPE
    HPCJTA  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
    HPCJTB  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
    HPCK3A  TEAMTRYSAP PGDAPQPTYD LELITSCSSN VSVARDDKGK RYYYLTRDAT
 HPCPLYPRE  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDGAGK RVYYLTRDPT
   HPCPOLP  TEAMTRYSAP PGDPPRPEYD LELITSCSSN VSVALGPQGR RRYYLTRDPT
     HPCPP  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
  HPCUNKCD  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
     MKC1A  TEAMTRYSAP PGDPPQPEYD LELITSCSSN VSVAHDASGK RVYYLTRDPT
     NDM59  TEAMTRYSAP PGDPPRPEYD LELITSCSSN VSVAXGPQGR RRYYLTRDPT
      NZLI  TEAMTRYSAP PGDAPQATYD LELITSCSSN VSVARDDKGR RYYYLTRDAT
      SA13  TEAMTRYSAP PGDPPVPAYD LELVTSCSSN VSVAHDASGN RVYYLTRDPQ
     Th580  TDAMIRYSAP PGDAPQPTYD LELITSCSSN VSVAHDGTGQ RYYYLTRDCT
 Type_3a_CB TEAMTRYSAP PGDAPQPTYD LELITSCSSN VSVARDDKGR RYYYLTRDAT
   TypeV_D  TEAMTRYSAP PGDAPQPTYD LELITSCSSN VSVALDNKGK RYYYLTRDAT
     VN004  TEAMTRYSAT PGDEPSPTYD LELITSCSSN VSVAHDGDGR RYYYLTRDPV
     VN235  TEAMTRYSAP PGDDPQPEYD LELITSCSSN VSVAHDGAGQ RYYYLTRDPL
     VN405  TEAMTRYSAP PGDEPHPAYD LELITSCSSN VSVAHDHTGQ RYYYLTRDPT

            2851                                              2900
     BEBE1  TPLARAAWET ARHSPVNSWL GNIIQYAPTV WVRMVLMTHF FSVLMAQDTL
     D89815 TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
ED43type_4  TPLARAVWET VRHTPVNSWL GNIIVYAPTI WVRMILMTHF FSILQSQEAL
      HC_C2 TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
      HC_G9 TPLARAAWET ARHTPVNSWL GNIIMFAPTL WVRMVLMTHF FSILIAQEHL
   HCU16326 TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
  HCV_H_CMR TPLARAAWET ARHTPVNSWL GNIIMFAPTL WARMILMTHF FSVLIARDQL
     HCV_J1 TPLARAAWET ARHTPVNSWL GNIIMFAPTL WARMILMTHF FSVLIARDQL
   HCV_J483 IPLARAAWET ARHTPVNSWL GNIIMYAPAL WARMILMTHF FSILLAQEQL
     HCV_J8 TPITRAAWET VRHSPVNSWL GNIIQYAPTI WVRMVIMTHF FSILLAQDTL
    HCV_JK1 TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
     HCV_JS TPIARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
  HCV_K1_R1 TPLARAAWET AKSTPVNSWL GNIIMFAPTL WVRMILMTHF FSILLAQEQL
  HCV_K1_R2 TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
  HCV_K1_R3 TPLARAAWET ARHTPVNSWL GNIIMYGPTL WARMILMTHF FSNLLAQEYL
  HCV_K1_S1 TPLARAAWET ARSTPVNSWL GNIIMFAPTL WVRMILMTHF FSILLAQEQL
  HCV_K1_S2 TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
  HCV_K1_S3 TPLARAAWET ARHTPVNSWL GNIIMYGPTL WARMILMTHF FSNLLAQEYL
    HCV_L2  TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
     HCV_N  TPIARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
   HCV12083 TPLARAAWET ARHTPVNSWL GNIIMFAPTI WVRMVLMNHF FSILQSQEQL
    HCV1480 VPLAKAAWET AKHSPVNSWL GNIIMYAPTL WARIVLMTHF FSVLQSQEQL
```

```
HCVPOLYP  TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
    HD_1  TPLARAAWET ARHTSVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
 HPCCGAA  TPLARAAWET ARHTPVNSWL GNIIMFAPTL WARMILMTHF FSVLIARDQL
   HPCFG  TPLARAAWET ARHTPVNSWL GNIIMFAPTI WVRMVLITHF FSILQAQEQL
HPCGENANTI TPLARAAWAT ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
 HPCGENOM  IPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
  HPCHUMR  TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
    HPCJ  TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
   HPCJCG  TPLARAAWET VRHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
 HPCJK046  NVLARAAWET ARHTPVNSWL GNIIMYAPTI WVRMVLMTHF FGILQPQEQL
 HPCJK049  TPLARAAWET ARHTPVNSWL GNIIMYAPTI WVRMVIMTHF FSILQAQEQL
   HPCJTA  TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
   HPCJTB  TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
   HPCK3A  TPLARAAWET ARHTPVNSWL GSIIMYAPTI WVRMVMMTHF FSILQSQEIL
HPCPLYPRE  TPLARAAWET ARHTPVNSWL GNIIMFAPTL WARMILMTHF FSVLIARDQL
  HPCPOLP  TPIARAAWET VRHSPVNSWL GNIIQYAPTI WARMVLMTHF FSILMAQDTL
    HPCPP  TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
 HPCUNKCD  TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
   MKC1A  TPLARAAWET ARHTPVNSWL GNIIMYAPTL WARMILMTHF FSILLAQEQL
   NDM59  TPLSRAAWET VRHSPVNSWL GNIIQYAPTI WVRMVLMTHF FSILMAQDTL
    NZLI  TPLARAAWET ARHTPVNSWL GNIIMYAPTI WVRMVMMTHF FSILQSQEIL
    SA13  VPLARAAWET AKHSPVNSWL GNIIMYAPTL WARIVLMTHF FSVLQSQEQL
   Th580  TPLARAAWET ARHTPVNSWL GNIIMYAPTI WVRMVLMTHF FSILQCQEQL
Type_3a_CB TPLARAAWET ARHTPVNSWL GNIIMYAPTI WVRMVMMTHF FSILQSQEIL
  TypeV_D  TPLARAAWET ARHTPVNSWL GNIIMYAPTI WVRMVMMTHF FSILQSQEIL
    VN004  TPLARAAWET ARHTPVNSWL GNIIMYAPTI WVRMVLMTHF FQILQAQETL
    VN235  TPLSRAAWET ARHTPVNSWL GNIIMYAPTI WVRMVLMTHF FAILQSQEIL
    VN405  TPLSRAAWET ARHTPVNSWL GNIIMYAPAI WVRMVLMTHF FQILQAQEQL

           2901                                                2950
    BEBE1  DQDLNFEMYG AVYSVSPLDL PAIIERLHGL EAFSLHSYSP HELTRVAAAL
    D89815  EKALDCQIYG ATYSIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
ED43type_4  EKALDFDMYG VTYSITPLDL PAIIQRLHGL SAFTLHGYSP HELNRVAGAL
    HC_C2  EKALECQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
    HC_G9  EKALDCEIYG AVHSVQPLDL PEIIQRLHGL SAFSLHSYSP GEINRVAACL
  HCU16326  EKTLDCQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
 HCV_H_CMR  EQALNCEIYA ACYSIEPLDL PPIIQRLHGL SAFLLHSYSP GEVNRVAACL
    HCV_J1  EQALDCEIYG ACYSIEPLDL PPIIQRLHGL SAFSLHSYSP GEINRVAACL
  HCV_J483  EQALDCQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
    HCV_J8  NQNLNFEMYG AVYSVNPLDL PAIIERLHGL EAFSLHTYSP HELSRVAATL
   HCV_JK1  EKALGCQIYG ATYFIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
    HCV_JS  EKALDCQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
  HCV_K1_R1  EKALDCQIYG ACYSIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
  HCV_K1_R2  EKALDCQIYG ACYSIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
  HCV_K1_R3  DKALDCQIYE AIYSIGPLDL PQVIQRLHGL SAFSLHSYSP GEINRVASCL
  HCV_K1_S1  EKALDCQIYG ACYSIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
  HCV_K1_S2  EKALDCQIYG ACYSIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
  HCV_K1_S3  DKALDCQIYE AIYSIGPLDL PQVIQRLHGL SAFSLHSYSP GEINRVASCL
    HCV_L2  EKALECQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
     HCV_N  EKALDCQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
  HCV12083  EKAFDFDIYG VTYSVSPLDL PAIIQRLHGM AAFSLHGYSP VELNRVGACL
   HCV1480  EKTLAFEMYG SVYSVTPLDL PAIIQRLHGL SAFSLHSYSP SEINRVASCL
 HCVPOLYP  EKALDCQIYG ACYSIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
    HD_1  EKALDCQIYG ACYSIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
 HPCCGAA  EQALNCEIYG ACYSIEPLDL PPIIQRLHGL SAFSLHSYSP GEINRVAACL
   HPCFG  ERALDFEMYG ATYSVTPLDL PAIIERLHGL SAFSLHGYSP TELNRVAGAL
HPCGENANTI  EKALDCQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
 HPCGENOM  EKALDCQIYG AYYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
  HPCHUMR  EKALDCQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
    HPCJ  GKALDCQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
   HPCJCG  EKALDCQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
```

```
HPCJK046   HKALDFDMYG VTYNITPLDL PQIIQRLHGM AAFSLHGYSP GELNRVGACL
HPCJK049   EKALDFEMYG AVYSVTPLDL PAIIERLHGL SAFSLHSYSP VELNRVAGAL
  HPCJTA   EKALDCQIYG ACYSIEPLDL PQIIQRLHGL SAFSLHSYSP VELNRVAGAL
  HPCJTB   EKALDCQIYG ACYSIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
  HPCK3A   DRPLDFEMYG ATYSVTPLDL PAIIERLHGL SAFSLHSYSP GEINRVASSL
HPCPLYPRE  EQALDCEIYG ACYSIEPLDL PPIIQRLHGL SAFSVHSYSP VELNRVAGTL
  HPCPOLP  DQNLNFEMYG AVYSVSPLDL PAIIERLHGL SAFSLHSYSP GEINRVAACL
    HPCPP  EKALDCQIYG ATYSIEPLDL PQIIQRLHGL DAFSLHTYTP HELTRVASAL
 HPCUNKCD  EKTLDCQIYG ACYSIEPLDL PQIIERLHGL SAFSLHSYSP GEINRVASCL
   MKC1A   EKALDCQIYG ATYSIEPLDL PQIIQRLHGL SAFSLHSYSP GEINRVASCL
   NDM59   DQNLNFEMYG XVYSVSPLDL PAIIERLHGL SAFSLHSYSP GEINRVASCL
    NZLI   DRPLDFEMYG ATYSVTPLDL PAIIERLHGL DAFSLHTYSP HELTRVASAL
    SA13   EKALAFEMYG SVYSVTPLDL PAIIQRLHGL SAFTLHSYSP VELNRVAGTL
   Th580   EAALNFDMYG VTYSVTPLDL PAIIQRLHGL SAFTLHSYSP SEINRVSSCL
Type_3a_CB DRPLDFEMYG ATYSVTPLDL PAIIQRLHGM AAFSLHGYSP TELNRVGASL
  TypeV_D  DRPLDFEMYG ATYSVTPLDL PAIIERLHGL SAFTLHSYSP VELNRVAGTL
   VN004   DRALDFDIYG VTYSITPLDL PAIIERLHGL SAFTLHSYSP VELNRVAGTL
   VN235   HKALDFDMYG VTYSVTPLDL PVIIQRLHGM AAFSLHGYSP DELNRVASCL
   VN405   DKVLDFDMYG VTYSVSPLQL PAIIQRLHGM AAFSLHGYSP TELNRVGACL
```

```
           2951                                                 3000
    BEBE1  RKLGAPPLRA WKSRARAVRA SLISRGGSAA TCGRYLFNWA VRTKLKLTPL
   D89815  RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VKTKLKLTPI
ED43type_4 RKLGVPPLRA WRHRARAVRA KLIAQGGRAK ICGIYLFNWA VKTKLKLTPL
    HC_C2  RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
    HC_G9  RKLGVPPLRA WRHRARSVRA TLLSQGGRAA ICGKYLFNWA VKTKLKLTPL
  HCU16326 RKLGVPPLRA WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
 HCV_H_CMR RKLGVPPLRA WRHRARSVRA RLLSRGGRAA ICGKYLFNWA VRTKLKLTPI
   HCV_J1  RKLGVPPLRA WRHRARSVRA RLLSRGGRAA ICGKYLFNWA VRTKLKLTPI
  HCV_J483 RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGRYLFNWA VRTKLKLTPI
   HCV_J8  RKLGAPPLRA WKSRARAVRA SLIAQGARAA ICGRYLFNWA VRTKLKLTPI
  HCV_JK1  RKLGVPPLRV WRHRARSVRA KLLSQGGRAA ICGKYLFNWA VKTKLKLTPL
   HCV_JS  RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
 HCV_K1_R1 RKLGVPPLRV WRHRARGVRA KLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
 HCV_K1_R2 RKLGVPPLRT WRHRARSVRA KLLSQGGRAA TCGRYLFNWA VKTKLKLTPI
 HCV_K1_R3 RKLGVPPLRV WRHRARSVRA KLLSQGGRAA ICGKYLFNWA VRTKLKLTPI
 HCV_K1_S1 RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VKTKLKLTPI
 HCV_K1_S2 RKLGVPPLRT WRHRARSVRA KLLSQGGRAA TCGRYLFNWA VKTKLKLTPI
 HCV_K1_S3 RKLGVPPLRV WRHRARSVRA KLLSQGGRAA ICGKYLFNWA VRTKLKLTPI
   HCV_L2  RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
    HCV_N  RKLGVPPLRV WRHRARRVRA KLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
  HCV12083 RKLGVPPLRV WRHRARNVRA KLLSQGGRAA TCGKYLFNWA VKTKLKLTPI
  HCV1480  RKLGVLPSRA WRHRARAVRA KLIAQGGKAA ICGKYLFNWA VKTKLKLTPL
 HCVPOLYP  RKLGVPPLRA WRHRARAVRA KLIAQGGRAA ICGIYLFNWA VKTKRKLTPL
    HD_1   RKLGVPPLRA WRHRARSVRA KLLSQGGRAA TCGRYLFNWA VKTKLKLTPI
  HPCCGAA  RKLGVPPLRV WRHRARSVRA KLLSPGGEGS TCGKYLFNWA VRTKLKLTPI
   HPCFG   RKLGVPPLRA WRHRAWSVRA RLLARGGKAA ICGKYLFNWA VRTKLKLTPI
HPCGENANTI RKLGIPPLRA WRHRARAVRA KLIAQGGKAR ICGLYLFNWA VRTKTKLTPL
 HPCGENOM  RKLGVPPLRA WRHRARSVRA KLLSQGGRAA TCGRYLFNWA VKTKLKLTPI
  HPCHUMR  RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VKTKLKLTPI
    HPCJ   RKLGVPPLRV WRHRARSVRA RLLSQGGRAA TCGKYLFNWA VKTKLKLTPI
   HPCJCG  RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
  HPCJK046 RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VKTKLKLTPI
  HPCJK049 RKLGAPPLRA WRHRARAVRA KLIAQGGKAA ICGMYLFNWA VKTKLKLTPL
   HPCJTA  RKLGIPPLRA WRHRARAVRA KLISQGGKAK ICGLYLFNWA VRTKAKLTPL
   HPCJTB  RKLGVPPLRV WRHRARSVRA RLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
   HPCK3A  RKLGVPPLRV WRHRARSVRA RLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
  HPCK3A   RKLGCPPLRA WRHRARAVRA KLIAQGGRAK ICGLYLFNWA VRTKTKLTPL
HPCPLYPRE  RKLGVPPLRA WRHRARSVRA RLLARGGRAA ICGKYLFNWA VRTKLKLTPI
  HPCPOLP  RKLGAPPLRA WKSRARAVRA SLISRGGRAA VCGRYLFNWA VKTKLKLTPL
    HPCPP  RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VKTKLKLTPI
 HPCUNKCD  RKLGVPPLRA WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VRTKLKLTPI
```

```
      MKC1A  RKLGVPPLRV WRHRARSVRA KLLSQGGRAA TCGKYLFNWA VKTKLKLTPI
      NDM59  RKLGAPPLRA WKSRARAVRA SLISRGGRAA ICGRYLFNWA VKTKLKLTPL
       NZLI  RKLGCPPLRA WRHRARAVRA KLIAQGGKAK ICGLYLFNWA VRTKTNLTPL
       SA13  RKLGVPPLRA WRHRARAVRA KLIAQGGKAA ICGIYLFNWA VKTKRKLTPL
      Th580  RKLGAPPLRA WRHRARAVRA KLIAQGGKAA ICGKYLFNWA VKTKLKLTPL
  Type_3a_CB RKLGCPPLRA WRHRARAGRA KLIAQGGKAK ICGLYLFNWA VRTKTKLTPL
    TypeV_D  RKLGCPPLRA WRHRARAVRA KLIAQGGKAK ICGLYLFNWA VRTKTNLTPL
      VN004  RKLGAPPLRA WRHRARAVRA KLIAQGGKAA VCGKYLFNWA IKTKLRLTPL
      VN235  RKLGAPPLRA WRHRARAVRA KLIAQGGKHA ICGKYLFNWA VRTKLKLTPL
      VN405  RKLGAPPLRA WRHRARAVRA KLIAQGGGAA ICGKYLFNWA VKTKLKLTPI

             3001                                                3050
      BEBE1  PAARLLDLSS WFTVSAGGGD IYHSVSRARP RLLLLGLLLL CVGVGIFLLP
      D89815 PEASQLDLSG WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
  ED43type_4 PAAAKLDLSG WFTVGAGGGD IYHSMSHARP RYLLLCLLIL TVGVGIFLLP
       HC_C2 PAASRLDLSG WFVAGYGGGD IYHSLSRARP RWFMLCLLLL SVGVGIYLLP
       HC_G9 PSASQLDLSN WFTGGYSGGD IYHSVSHVRP RWFFWCLLLL SVGVGIYLLP
    HCU16326 PAASRLDLSG WFVAGYSGGD IYHSLSRARP RWFMLCLLLL SVGVGIYLLP
   HCV_H_CMR AAAGRLDLSG WFTAGYSGGD IYHSVSHARP RWFWFCLLLL AAGVGIYLLP
      HCV_J1 AAAGRLDLSG WFTAGYSGGD IYHSVSHARP RWFWFCLLLL AAGVGIYLLP
   HCV_J483  PAASQLDLSG WFVAGYSGGD IYHSLSRARP RWFLLCLLLL SVGVGIYLLP
      HCV_J8 PEASRLDLSG WFTVGAGGGD IYHSVSHARP RLLLLCLLLL SVGVGIFLLP
     HCV_JK1 PAASQLDLSG WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
      HCV_JS PAASRLDLSG WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
   HCV_K1_R1 PAASQLDLSN WFVAGYSGGD VYHSLSRARP RWFMLCLLLL SVGVGIYLLP
   HCV_K1_R2 PAASQLDLSG WFVAGYSGGD IYHSVSRARP RWFMWCLLLL SVGVGIYLLP
   HCV_K1_R3 PAASQLDLSS WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
   HCV_K1_S1 PAASQLDLSN WFVAGYSGGD VYHSLSRARP RWFMLCLLLL SVGVGIYLLP
   HCV_K1_S2 PAASQLDLSG WFVAGYSGGD IYHSVSRARP RWFMWCLLLL SVGVGIYLLP
   HCV_K1_S3 PAASQLDLSS WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
      HCV_L2 PAASRLDLSS WFVAGYSGGD IYHSVSHARP RWFMLCLLLL SVGVGIYLLP
       HCV_N PAASQLDLSG WFVAGYSGGD IYHSLSRARP RWFMLCLLLL SVGVGIYLLP
    HCV12083 VSASKLDLSG WFVAGYDGGD IYHSVSQARP RFLLLGLLLL TVGVGIFLLP
     HCV1480 ADADRLDLSS WFTVGAGGGD IYHSMSRARP RNLLLCLLLL SVGVGIFLLP
    HCVPOLYP PAASQLDLSN WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
        HD_1 PAAFQLDLSG WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
     HPCCGAA TAAGRLDLSG WFTAGYSGGD IYHSVSHARP RWFWFCLLLL AAGVGIYLLP
       HPCFG PTAGQLDLSS WFTVGVGGND IYHSVSRART RHLLLCLLLL TVGVGIFLLP
  HPCGENANTI PAASQLDLSK WFVAGYGGGD IYHSLSRARP RWFMLCLLLL SVGVGIYLLP
    HPCGENOM PAASRLDLSG WFVAGYSGGD IYHSLSRARP RWFMLCLLLL SVGVGIYLLP
     HPCHUMR PAASRLDLSG WFVAGYSGGD IYHSLSRARP RWFMLCLLLL SVGVGIYLLP
        HPCJ PAASQLDLSS WFVAGYSGGD IYHSLSRARP RWFMLCLLLL SVGVGVYLLP
     HPCJCG PAASQLDLSG WFVAGYNGGD IYHSLSRARP RWFMLCLLLL SVGVGIYLLP
    HPCJK046 RDAHRLDLSG WFVAGYSGGD IFHSVSHARP RVLLLCLLLL TVGVGIFFLP
    HPCJK049 PQAGLLDLSR WFTVGAGGND IYHSVSRARS RHLLLGLLLL TVGVGIFLLP
     HPCJTA PAASQLDLSS WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
     HPCJTB PAASQLDLSS WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
     HPCK3A PAAGQLDLSS WFTVGVGGND IYHSVSRART RYLLLCLLLL TVGVGIFLLP
   HPCPLYPRE AAAGQLDLSG WFTAGYSGGD IYHSVSHARP RWIWFCLLLL AAGVGIYLLP
     HPCPOLP PEARLLDLSS WFTVGAGGGD IYHSVSRARP RLLLLGLLLL FVGVGLFLLP
       HPCPP PEASQLDLSG WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
    HPCUNKCD PAASRLDLSG WFVAGYSGGD IYHSLSRARP RWFMLCLLLL SVGVGIYLLP
      MKC1A PEASQLDLSG WFVAGYSGGD IYHSLSRARP RWFMWCLLLL SVGVGIYLLP
      NDM59 PEARLLDLSS WFTVGAGGGD IYHSVSRARP RLLLLSLLLL LVGVGLFLLP
       NZLI PAAGQLDLSS WFTVGVGGND IYHSVSRART RHLLLCLLLL TVGVGIFLLP
       SA13 ADADRLDLSS WFTVGAGGGD IYHSMSRARP RCILLCLLLL TVGVGIFLLP
      Th580 AAASQLDLSG WFVAGYDGGD IYHSVSRARP RLLLLGLLLL TVGVGIFLLP
  Type_3a_CB PRAGQLDLSI WFTVGVGGND IYHSVSRART RYLLLCLLLL TVGVGIFLLP
    TypeV_D  PATGQLDLSS WFTVGVGGND IYHSVSRART RYLLLCLLLL TVGVGIFLLP
      VN004  RGASALDLSG WFTSGYGGGD VYHSASRARP RFLLLCLLLL SVGVGIFLLP
      VN235  RGAANLDLSG WFVSGGSGGD IFHSVSRARP RNLLLCLLLL TVGVGIFLLP
```

```
     VN405  PDAARLDLSG WFISGFSGGD IYHSVSRARP RIFLLCLLLL SVGVGIFLLP

              3051
      BEBE1  AR
     D89815  NR
ED43type_4   AR
      HC_C2  NR
      HC_G9  NR
   HCU16326  NR
  HCV_H_CMR  NR
     HCV_J1  NR
   HCV_J483  NR
     HCV_J8  AR
    HCV_JK1  NR
     HCV_JS  NR
  HCV_K1_R1  NR
  HCV_K1_R2  NR
  HCV_K1_R3  NR
  HCV_K1_S1  NR
  HCV_K1_S2  NR
  HCV_K1_S3  NR
     HCV_L2  NR
      HCV_N  N.
   HCV12083  AR
    HCV1480  AR
   HCVPOLYP  NR
       HD_1  NR
    HPCCGAA  NR
      HPCFG  AR
HPCGENANTI   NR
   HPCGENOM  NR
    HPCHUMR  NR
       HPCJ  NR
     HPCJCG  NR
   HPCJK046  PR
   HPCJK049  AR
     HPCJTA  NR
     HPCJTB  NR
     HPCK3A  AR
 HPCPLYPRE   NR
    HPCPOLP  AR
      HPCPP  NR
   HPCUNKCD  NR
      MKC1A  NR
      NDM59  AR
       NZLI  AR
       SA13  AR
      Th580  AR
Type_3a_CB   AR
   TypeV_D   AR
      VN004  AR
      VN235  AR
      VN405  AR
```

Table 23. HIV Fusion Construct

EP-HIV-1090
MGMQVQIQSLFLLLLWVPGSRGKLVGKLNWAGAAILKEPVHGVNAACPKVSFEPIKIPIHYCAPA
KAKFVAAWTLKAAAKAEPVRPQVPLGAAKLTFLCVTLGAAAVLAEAMSQVKVYLAWVPAHKG
AAAAIFQSSMTKKTTLFCASDAKNIPYNFQSQGVVKHPVHAGPIANVTVYYGVPVWKKAAAQMA
VFIHNFKNAAAYPLASLRSLFNLTFGWCFKLNRILQQLLFINAKIQNFRVYYRKAAVTIKIGGQLKK
VPLQLPPLKAMTNNPPIPV

ATGGGAATGCAGGTGCAGATCCAGAGCCTGTTTCTGCTCCTCCTGTGGGTGCCCGGATCCAGA
GGAAAGCTGGTGGGCAAACTCAACTGGGCCGGAGCTGCAATCCTGAAGGAGCCCGTCCACGG
GGTGAATGCCGCTTGCCCTAAAGTCAGCTTCGAACCAATTAAGATCCCCATTCATTACTGTGC
ACCTGCCAAAGCTAAGTTTGTGGCCGCTTGGACCCTCAAGGCCGCTGCAAAAGCCTTCCCAGT
GAGGCCCCAGGTGCCTCTGGGCGCCGCTAAACTCACACCACTGTGCGTCACTCTGGGAGCCGC
TGCAGTGCTGGCAGAGGCCATGTCCCAAGTGAAGGTGTATCTGGCTTGGGTGCCCGCCCACAA
GGGGGCCGCTGCAGCCATCTTTCAGTCTAGCATGACCAAGAAAACAACTCTGTTCTGTGCCTC
CGACGCTAAGAACATCCCTTATAATCCACAGTCTCAGGGCGTGGTCAAGCATCCCGTGCACGC
CGGACCTATTGCTAACGTGACCGTGTACTATGGGGTCCCAGTGTGGAAGAAAGCCGCTGCACA
GATGGCCGTGTTTATTCACAATTTCAAAAACGCCGCTGCATACCCCCTCGCCAGCCTGAGATC
CCTCTTCAACCTGACATTCGGCTGGTGCTTTAAGCTGAACCGGATCCTGCAGCAACTGCTCTTT
ATCAATGCTAAAATCCAGAACTTCCGCGTCTACTATAGGAAGGCTGCAGTGACTATCAAAATT
GGCGGACAACTGAAGAAAGTGCCTCTCCAGCTGCCCCCTCTCAAGGCAATGACCAACAATCC
CCCTATCCCAGTCTGA

Table 24. HBV GCR-3697 Fusion Construct

GCR-3697 SEQ ID NO:__	**Polynucleotide** 1 Start ↑* * * ATGGGCATGCAGGTGCAGATCCAGAGCCTGTTCCTGCTCCTGCTGTGGGTGCCAGGAAGCAGAGGCTTTCTC CTGTCCCTGGGCATCCACCTGAACGCCGCTGCAAAGTACACCAGCTTCCCCTGGCTGCTCAACGCCGCTGCC CGGTTCAGCTGGCTGTCCCTGCTCGTGCCCTTCAACGCAGCCTTCCCCCACTGCCTGGCCTTCAGCTACATGA AAGCAGCCCTGGTGGTCGACTTCTCCCAGTTCAGCCGGGGAGCCATCCTGCTCCTGTGCCTGATCTTTCTGCT CAACGCCGCTGCCCACACCCTGTGGAAGGCTGGCATCCTGTACAAGAAAGCCTGGATGATGTGGTACTGGG GACCCAGCCTGTACAAGGCATATCCAGCCCTGATGCCCCTGTACGCCTGCATCGGAGCTGCCGCATGGCTGA GCCTCCTGGTGCCCTTCGTGAACGCCGCTGCCGGGTTCCTGCTGACAAGAATCCTGACCATCAACGCCGCAG CCATTCCTATCCCCTCCAGCTGGGCCTTCAAGGCAGCCGCCGAGTACCTGGTGAGCTTCGGAGTCTGGAACC TGCCCAGCGACTTCTTTCCCAGCGTGAAAGCCGCAGCCTTCCTGCCCTCCGACTTCTTTCCCAGCGTGAAGGC CGCAGCCGATCTCCTGGACACCGCTAGCGCCCTGTACAACAGCTGGCCCAAGTTCGCCGTGCCCAACCTGAA GGCCGCAGCCAGCGCCATCTGCAGCGTGGTCAGACGGAAGCTGTCCCTCGATGTGAGCGCCGCTTTCTACAA CGCCGCCGCAAAGTTCGTGGCCGCCTGGACCCTGAAAGCCGCTGCCAAGGCAGCCAACGTGAGCATCCCCT GGACCCACAAAGGAGCCGCAGGACTGAGCCGGTATGTGGCCAGACTGAACGCCGCTGCCAGCACCCTGCCC GAGACCACAGTGGTCAGACGGAAGCACCCCGCCGCCATGCCCCACCTGCTGAAGGCCGCAGCCCGGTGGAT GTGCCTCAGACGGTTCATCATCAACGCTTCCTTCTGTGGCAGCCCCTACAAGGCCGCCTACATGGATGACGT GGTCCTGGGAGTGAACGCCCTCTGGTTCCACATCAGCTGCCTGACCTTCAAAGCCGCTGCCACACCCGCAAG AGTGACCGGAGGCGTGTTCAAGGCTGCAGCCCTGACCTTCGGCCGGGAGACCGTGCTGGAGTACAAGCAGG CCTTCACCTTCAGCCCCACCTACAAGAACGCCGGCACCAGCTTTGTGTACGTCCCAAGCGCCCTGAATCCCG CAGACGGCCCCGGCCCCGGACTGTGCCAGGTGTTCGCCGATGCCACACCAACCGGATGGGGCCTGGGCCCT GGACCCGGCAGACACTACCTGCATACCCTGTGGAAGGCAGGAATCCTGTACAAAGGCCCCGGCCCTGGACC CCATCACACCGCTCTGCGGCAGGCCATCCTGTGCTGGGGCGAGCTCATGACTCTGGCAGGACCCGGCCCCGG CGAATCCAGGCTGGTGGTGGACTTTAGCCAGTTCTCCAGAGGCAACGGACCCGGCCCAGGACCCTTCCTGCT CGCCCAGTTCACCAGCGCCATCTGCAGCGTGGTCGGACCTGGCCCAGGACTGGTGCCCTTCGTGCAGTGGTT CGTCGGCCTCAGCCCCACCGTCGGACCTGGCCCCGGCCTGCACCTCTACAGCCACCCTATCATTCTGGGCTT CAGAAAGATCGGACCAGGCCCCGGCTCCAGCAACCTGTCCTGGCTCAGCCTGGACGTCAGCGCAGCCTTCG GACCCGGCCCTGGCCTGCAGAGCCTGACCAACCTGCTCAGCAGCAACCTCAGCTGGCTGGGCCCAGGACCC GGCGCAGGCTTCTTTCTGCTCACCAGAATCCTGACCATCCCTCAGAGCGGCCCCGGACCAGGCGTGAGCTTC GGCGTGTGGATTCGGACTCCTCCCGCCTACAGACCCCCAAATGCCCCCATCGGCCCAGGACCCGGCGTCGGA CCTCTGACTGTGAACGAGAAGCGGAGACTGAAGCTGATCGGCCCCGGACCAGGCAAACAGTGCTTCAGGAA GCTCCCTGTGAACAGACCTATCGACTGGGGCCCCGGACCCGGCGCAGCCAACTGGATTCTGAGAGGCACCA GCTTCGTGTACGTCCCTGGACCCGGCCCTGGCAAGCAAGCCTTCACCTTCAGCCCCACCTACAAGGCATTCC TGTGCGGATAG ↓ Stop 2232
GCR-3697 SEQ ID NO:__	***Polypeptide*** 1 ↑ MGMQVQIQSLFLLLLWVPGSRGFLLSLGIHLNAAAKYTSFPWLLNAAARFSWLSLLVPFNAAFPHCLAFSYMKA ALVVDFSQFSRGAILLLCLIFLLNAAAHTLWKAGILYKKAWMMWYWGPSLYKAYPALMPLYACIGAAAWLSLL VPFVNAAAGFLLTRILTINAAAIPIPSSWAFKAAAEYLVSFGVWNLPSDFFPSVKAAAFLPSDFFPSVKAAADLLD TASALYNSWPKFAVPNLKAAASAICSVVRRKLSLDVSAAFYNAAAKFVAAWTLKAAAKAANVSIPWTHKGAA GLSRYVARLNAAASTLPETTVVRRKHPAAMPHLLKAAARWMCLRRFIINASFCGSPYKAAYMDDVVLGVNAL WFHISCLTFKAAATPARVTGGVFKAAALTFGRETVLEYKQAFTFSPTYKNAGTSFVYVPSALNPADGPGPGLCQ VFADATPTGWGLGPGPGRHYLHTLWKAGILYKGPGPGPHHTALRQAILCWGELMTLAGPGPGESRLVVDFSQFS RGNGPGPGPFLLAQFTSAICSVVGPGPGLVPFVQWFVGLSPTVGPGPGLHLYSHPIILGFRKIGPGPGSSNLSWLSL DVSAAFGPGPGLQSLTNLLSSNLSWLGPGPGAGFFLLTRILTIPQSGPGPGVSFGVWIRTPPAYRPPNAPIGPGPGV GPLTVNEKRRLKLIGPGPGKQCFRKLPVNRPIDWGPGPGAANWILRGTSFVYVPGPGPGKQAFTFSPTYKAFLCG ↓ 744

Table 25. HBV AOSIb2 Fusion Construct

HBV AOSIb2 SEQ ID NO:__	**Polynucleotide** 1 Start ↑ ATGGGAATGCAGGTGCAGATCCAGAGCCTGTTTCTGCTCCTCCTGTGGGTGCCCGGGTCCAGAGGACACAC CCTGTGGAAGGCCGGAATCCTGTATAAGGCCAAGTTCGTGGCTGCCTGGACCCTGAAGGCTGCCGCTTTCCT GCCTAGCGATTTCTTTCCTAGCGTGAACTTCCTGCTGTCCCTGGGAATCCACCTGTATATGGATGACGTGGTG CTGGGAGTGGGACTGTCCAGGTACGTGGCTAGGCTGTTCCTGCTGACCAGAATCCTGACCATCTCCACCCTG CCAGAGACCACCGTGGTGAGGAGGCAGGCCTTCACCTTTAGCCCTACCTATAAGGGAGCCGCTGCCTGGCT GAGCCTGCTGGTGCCCTTTGTGAATATCCCTATCCCTAGCTCCTGGGCTTTCAAGACCCCAGCCAGGGTGAC CGGAGGAGTGTTTAAGGTGGGAAACTTCACCGGCCTGTATAACCTGCCCAGCGATTTCTTTCCTAGCGTGAA GACCCTGTGGAAGGCCGGAATCCTGTACAAGAATGTGTCCATCCCTTGGACCCACAAGGGAGCCGCTCTGG TGGTGGACTTTTCCCAGTTCAGCAGAAATTCCGCTATCTGCTCCGTGGTGAGGAGAGCTCTGATGCCACTGT ATGCCTGTATCTGA ↓ Stop 657
HBV AOSIb2 SEQ ID NO:__	***Polypeptide*** 1 ↑ MGMQVQIQSLFLLLLWVPGSRGHTLWKAGILYKAKFVAAWTLKAAAFLPSDFFPSVNFLLSLGIHLYMDDVVL GVGLSRYVARLFLLTRILTISTLPETTVVRRQAFTFSPTYKGAAAWLSLLVPFVNIPIPSSWAFKTPARVTGGVFKV GNFTGLYNLPSDFFPSVKTLWKAGILYKNVSIPWTHKGAALVVDFSQFSRNSAICSVVRRALMPLYACI ↓ 219

Table 26. HCV Fusion Construct

HCV 4312(1P)
MGMQVQIQSLFLLLLWVPGSRGRLGVRATRKKAAAKTSERSQPRNLPGCSFSIFNDLMGYIPLVK
YLLPRRGPRLNTLCGFADLMGYRMYVGGVEHRKLLFNILGGWVKAAALADGGCSGGAYRLIVFP
DLGVKFWAKHMWNFISGIFWAKHMWNFKKAAAVLVGGVLAA
KKKYLVTRHADVLGFGAYMSKCTCGSSDLYHMWNFISGIFWAKHMWNFKKAAAVLVGGVLAA
AFLLLADARVLSAFSLHSYILAGYGAGVWMNRLIAFANAAAKFVAAWTLKAAA*

GAATTCGCCGCCACCATGGGAATGCAGGTGCAGATCCAAAGCCTGTTTCTGCTCCTCCTGTGG
GTGCCCGGCTCCAGAGGAAGGCTGGGCGTGAGAGCCACCCGGAAGAAGGCTGCCGCTAAAAC
AAGCGAGCGCTCCCAGCCCAGGAACCTGCCTGGATGCTCTTTCAGCATCTTTAATGACCTCAT
GGGGTACATTCCACTGGTGAAGTATCTGCTCCCCAGACGGGGCCCTCGCCTGAACACTCTCTG
TGGATTTGCTGATCTGATGGGGTACAGGATGTATGTCGGCGGAGTCGAACACAGAAAACTGCT
CTTCAACATCCTGGGCGGATGGGTGAAGGCTGCCGCTCTGGCCGACGGGGGATGCAGCGGCG
GAGCTTACAGGCTCATTGTCTTTCCCGATCTCGGAGTCAAATTTTGGGCAAAGCACATGTGGA
ATTTCATCGGGGTGGCCGGAGCCCTGGTCGCTTTTAAAAAAGCAGCTCTTCACCTTCTCCCCAA
GACGGAACCGATACCTCGTCGCCTACCAGGCCACTGTGGCTGCAGCTCTGCTCTTCCTGCTCC
TGGCCGATGCACTCATCTTCTGCCATTCCAAGAAAAAGTATCTGGTCACCAGACATGCTGACG
TGCTGGGGTTTGGCGCCTACATGAGCAAGTGCACCTGTGGCAGCTCCGACCTGTATCACATGT
GGAACTTTATTTCTGGAATCTTTTGGGCCAAGCACATGTGGAATTTTAAGAAAGCCGCTGCAG
TCCTGGTGGGCGGCGTCCTGGCAGCCGCTTTCCTGCTCCTGGCAGACGCCAGGGTGCTGTCTG
CCTTCAGCCTCCACTCCTACATCCTCGCAGGGTATGGCGCAGGCGTGTGGATGAATCGGCTGA
TCGCCTTTGCCAATGCTGCAGCTAAATTCGTGGCAGCCTGGACACTGAAAGCAGCTGCATGAG
GATCC

Table 27. ***Plasmodium falciparum*** **Fusion Construct**

P33
MGMQVQIQSLFLLLLWVPGSRGFMKAVCVEVNVTCGNGIQVRKGLIMVLSFLNAALFHIFDGDN
EIKAALLACAGLAYKKSFLFVEALFNAAPSDGKCNLYKAAQTNFKSLLRNLPSENERGYKAAGVS
ENIFLKNAAAYFILVNLLIKAAAILSVSSFLFVNTPYAGEPAPFKAAAKYKLATSVLKAAVFLIFFDL
FLNYYIPHQSSLKAAGLLGNVSTVGAVLLGGVGLVLNLACAGLAYKKAKFIKSLFHIFKAAFYFIL
VNLLKAFLIFFDLFLVKALFFIIFNKNYYGKQENWYSLKFVEALFQEYNAAAKFVAAWTLKAAAK
ILSVFFLANAVLAGLLGNVNFQDEENIGIYKAAALYISFYFIKAFILVNLLIFHNAALPYGRTNLKAA
HVLSHNSYEKNAAAKYLVIVFLI

GCCGCCACCATGGGAATGCAGGTGCAGATCCAGAGCCTGTTTCTGCTCCTCCTGTGGGTGCCC
GGATCCAGAGGATTTATGAAAGCTGTCTGTGTAGAGGTGAATGTAACATGCGGTAACGGAAT
TCAGGTGAGAAAGGGACTCATCATGGTACTCAGCTTTCTGAACGCAGCCCTGTTCCACATCTT
TGACGGAGACAATGAAATCAAAGCCGCATTGCTCGCCTGTGCCGGACTAGCCTATAAAAAGA
GTTTCCTTTTCGTTGAAGCACTATTTAAGGCAGCACCCAGTGACGGTAAATGCAACCTATATA
AAGCAGCTCAGACTAATTTCAAAAGCCTGTTAAGAAATCTGCCCTCAGAGAATGAAAGGGGT
TACAAAGCCGCCGGCGTGTCCGAGAATATTTTCCTGAAGAACGCCGCTGCTTATTTTATACTC
GTGAATCTACTCATAAAGGCAGCCGCAATCCTTTCAGTGTCCAGCTTTCTGTTTGTTAACACAC
CATATGCGGGCGAGCCGGCTCCTTTCAAGGCTGCAGCAAAATACAAGCTTGCCACATCAGTAT
TGAAAGCAGCTGTGTTTTTGATATTCTTTGATCTTTTTTTAAACTACTACATACCTCATCAGTCT
AGTCTTAAAGCAGCCGGGCTACTGGGGAACGTCTCTACTGTGGGGGCCGTCTTACTTGGAGGA
GTTGGCCTCGTGTTGAACCTCGCGTGCGCAGGTCTGGCCTACAAAAAAGCGAAATTCATCAAG
TCTCTGTTCCACATTTTTAAAGCCGCATTCTATTTCATACTAGTGAACCTTCTCAAAGCTTTCCT
GATCTTCTTCGATCTATTCCTCGTAAAAGCGCTATTCTTCATTATCTTTAACAAAAATTATTAC
GGCAAGCAAGAAAATTGGTACTCACTCAAGTTTGTAGAAGCTCTGTTCCAGGAATACAACGCC
GCTGCTAAATTCGTTGCAGCTTGGACCCTGAAAGCAGCTGCAAAGATCCTATCGGTCTTCTTTC
TCGCTAATGCCGTATTAGCAGGACTTCTAGGCAACGTGAACTTTCAAGACGAAGAGAATATAG
GCATCTACAAAGCCGCAGCACTGTACATTTCATTCTACTTCATCAAGGCCTTCATACTGGTCAA
CCTTCTGATATTTCATAATGCAGCACTGCCATATGGGAGAACCAACTTGAAAGCGGCCCACGT
GTTGAGCCACAACTCCTACGAGAAGAACGCCGCCGCGAAATATCTCGTCATTGTCTTCCTGAT
TTGA

Table 28. ***Mycobacterium tuberculosis*** **Fusion Construct**

TB1
MQVQIQSLFLLLLWVPGSRGRMSRVTTFTVKALVLLMLPVVNLMIGTAAAVVKALVLLMLPVGA
GLMTAVYLVGAAAMALLRLPVKRMFAANLGVNSLYFGGICVGRLPLVLPAVNAAAAKFVAAWT
LKAAAKAAARLMIGTAAAGFVVALIPLVNAMTYAAPLFVGAAAAMALLRLPLV

ATGCAGGTGCAGATCCAGAGCCTGTTTCTGCTCCTCCTGTGGGTGCCCGGATCCAGAGGAAGG
ATGAGCAGAGTGACCACATTCACTGTCAAGGCCCTGGTGCTCCTGATGCTCCCCGTCGTGAAC
CTGATGATCGGCACCGCTGCAGCCGTCGTGAAAGCTCTCGTCCTGCTCATGCTCCCTGTGGGA
GCAGGGCTGATGACAGCCGTGTACCTGGTCGGCGCTGCAGCCATGGCCCTCCTGCGGCTGCCA
GTGAAGCGCATGTTTGCTGCAAATCTGGGAGTCAACTCCCTCTATTTCGGGGGCATTTGCGTG
GGAAGGCTGCCCCTGGTGCTGCCTGCTGTGAATGCAGCCGCTGCCAAATTTGTCGCCGCTTGG
ACTCTGAAGGCAGCCGCTAAGGCCGCTGCAAGACTGATGATCGGGACCGCCGCTGCCGGCTT
CGTGGTCGCCCTGATTCCCCTGGTGAACGCCATGACATACGCAGCTCCTCTGTTTGTGGGAGC
CGCTGCAGCCATGGCTCTCCTGCGGCTGCCACTGGTGTGA

Table 29. Hepatitis B Virus Core Protein (SEQ ID NO: _)

MQLFHLCLIISCSCPTVQASKLCLGWLWGMDIDPYKEFGATVELLSFLPSDFFPSVRD
LLDTASALYREALESPEHCSPHHTALRQAILCWGELMTLATWVGVNLEDPASRDLVV
SYVNTNMGLKFRQLLWFHISCLTFGRETVIEYLVSFGVWIRTPPAYRPPNAPILSTLPE
TTVVRRRGRSPRRRTPSPRRRRSQSPRRRRSQSRESQC

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☑ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.
As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.